



DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/31, C07K 14/22, 16/12, A61K 39/095, C12Q 1/68, G01N 33/53		A2	(11) Numéro de publication internationale: WO 98/02547
			(43) Date de publication internationale: 22 janvier 1998 (22.01.98)
(21) Numéro de la demande internationale: PCT/FR97/01295		MERKER, Petra [DE/DE]; Cuvrystrasse 38, D-10997 Berlin (DE).	
(22) Date de dépôt international: 11 juillet 1997 (11.07.97)		(74) Mandataires: PEAUCHELLE, Chantal etc.; Cabinet Armengaud Ainé, 3, avenue Bugeaud, F-75116 Paris (FR).	
(30) Données relatives à la priorité: 96/08768 12 juillet 1996 (12.07.96)		FR	(81) Etats désignés: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, brevet ARIPO (GH, KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(71) Déposants (pour tous les Etats désignés sauf US): INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM) [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cedex 13 (FR). MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V., BERLIN [DE/DE]; Hofgartenstrasse 2, D-80539 Munich (DE). SMITHKLINE BEECHAM [GB/GB]; New Horizons Court, Brentford TW8 9EP (GB).		Publiée <i>Sans rapport de recherche internationale, sera republiée dès réception de ce rapport.</i>	
(72) Inventeurs; et			
(75) Inventeurs/Déposants (US seulement): NASSIF, Xavier [FR/FR]; 30, rue Labrouste, F-75015 Paris (FR). TINSELEY, Colin [FR/FR]; 156, rue de Vaugirard, F-75015 Paris (FR). ACHTMAN, Mark [DE/DE]; Neuenburgerstrasse 16, D-10969 Berlin (DE). RUELLE, Jean-Louis [BE/BE]; Résidence de la Lyre 18, B-1300 Limal (BE). VINALS, Carla [BE/BE]; Rue des Acacias 30, B-4000 Liège (BE).			
(54) Title: DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE <i>NEISSERIA MENINGITIDIS</i> SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL APPLICATIONS			
(54) Titre: ADN ET PROTEINES OU PEPTIDES SPECIFIQUES DES BACTERIES DE L'ESPECE <i>NEISSERIA MENINGITIDIS</i> , LEURS PROCEDES D'OBTENTION ET LEURS APPLICATIONS BIOLOGIQUES			
(57) Abstract			
<p>The DNA of the invention are characterised in that they concern the whole or part of genes, with their reading frame, to be found in <i>Neisseria meningitidis</i>, but not in <i>Neisseria gonorrhoeae</i>, or in <i>Neisseria lactamica</i> except the genes involved in the biosynthesis of the polysaccharide capsule, <i>frpA</i>, <i>frpC</i>, <i>opc</i>, <i>porA</i>, rotamase the sequence IC1106, IgA protease, pilin, pilC, transferrin binding proteins and opacity proteins. The invention also concerns the polypeptides corresponding to these DNA and the antibodies directed against these polypeptides. It is applicable in the prevention and the detection of meningococcus induced infections and meningitis.</p>			
(57) Abrégé			
<p>Les ADN de l'invention sont caractérisés en ce qu'il s'agit de tout ou partie de gènes, avec leur phase de lecture, présents chez <i>Neisseria meningitidis</i>, mais absents soit chez <i>Neisseria gonorrhoeae</i>, soit chez <i>Neisseria lactamica</i> à l'exception des gènes impliqués dans la biosynthèse de la capsule polysaccharidique, <i>frpA</i>, <i>frpC</i>, <i>opc</i>, <i>porA</i>, rotamase, de la séquence IC1106, des IgA protéases, de la piline, de pilC, des protéines qui lient la transferrine et des protéines d'opacité. L'invention vise également les polypeptides correspondant à ces ADN et les anticorps dirigés contre ces polypeptides. Applications à la prévention et à la détection d'infections à méningocoques et de méningites.</p>			

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Belarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LJ	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

ADN et protéines ou peptides spécifiques des bactéries de l'espèce *Neisseria meningitidis*, leurs procédés d'obtention et leurs applications biologiques.

5 L'invention est relative aux ADN, et aux protéines et peptides, spécifiques des bactéries de l'espèce *Neisseria meningitidis* (ci-après en abrégé Nm), à leur procédé d'obtention et à leurs applications biologiques, en particulier pour la prévention et la détection
10 d'infections à méningocoques et de méningites.

On sait que Nm constitue l'un des principaux agents de la méningite cérébrospinale.

Des études menées aux Etats-Unis ont montré que de 5 à 10% de la population sont porteurs asymptomatiques de
15 souche(s) de Nm. Les facteurs de transmission de Nm sont mal connus. Pour une proportion des personnes infectées, Nm pénètre le flux sanguin, où elle peut provoquer une méningococcémie et/ou progresse dans le flux cérébrospinal pour provoquer une méningite. Sans
20 traitement antibiotique rapide, l'infection peut se développer de manière fulgurante et devenir mortelle.

Comparée aux autres pathogènes, Nm présente la caractéristique de pouvoir franchir la barrière hémato-encéphalique afin de coloniser les méninges. L'étude de
25 la pathogénicité de Nm est donc non seulement importante dans le cadre de la méningite, mais aussi dans le cadre de toute maladie touchant le cerveau.

On conçoit alors l'intérêt de disposer d'outils spécifiques de cette espèce bactérienne pour les
30 applications envisagées ci-dessus.

Nm est génétiquement très proche des bactéries de l'espèce *Neisseria gonorrhoeae* (ci-après en abrégé Ng) et de l'espèce *Neisseria lactamica* (ci-après en abrégé Nl). Leur pathogénicité est toutefois très différente.

Nm colonise le nasopharynx, puis traverse l'épithélium pharyngé pour envahir l'espace sous-muqueux, étant alors responsable de septicémie et de méningite.

5 Ng est surtout responsable d'infections localisées du tractus génito-urinaire. Elle colonise la muqueuse génitale, puis traverse l'épithélium, envahit ensuite le sous-épithélium où elle se multiplie et est responsable d'une forte réaction inflammatoire. Des infections 10 gonococciques disséminées sont possibles, mais restent rares et sont le fait de seulement certaines souches. Quant à Nl, on considère qu'il s'agit d'une souche non pathogène, étant donné qu'elle n'est pas responsable d'invasion localisée ou générale.

15 Ainsi, une première considération amène à prendre en compte le fait que Nm et Ng, tout en étant des bactéries très proches, présentent des pouvoirs pathogènes différents.

20 Le génome de ces bactéries étant fortement homologue, seules des parties limitées du génome de Nm et de Ng doivent coder pour des facteurs de virulence spécifiques, responsables de leur pathogénèse.

25 Il est clair que Nm présente par rapport à Ng des séquences d'ADN qui lui sont spécifiques et qui doivent intervenir au niveau de l'expression de son pouvoir pathogène spécifique.

L'espèce Nm est subdivisée en sérogroupes basés sur la nature des polysaccharides capsulaires.

30 Au moins 13 sérogroupes ont été définis, parmi lesquels les sérogroupes A, B et C sont responsables d'environ 90% des cas de méningites. Les groupes A et C sont observés dans les formes épidémiques de la maladie. Le groupe B est le sérogroupe le plus couramment isolé en Europe et aux Etats-Unis.

La capsule, présente chez Nm et absente chez Ng, a servi de base pour l'élaboration de vaccins antiméningite méningococcique.

Les polysaccharides de la capsule de Nm ont été 5 utilisés pour l'élaboration d'un vaccin qui s'est montré efficace pour prévenir chez les adultes la méningite provoquée par les méningocoques de sérogroupes A, C, W135 et Y.

Cependant, le polysaccharide de Nm groupe C s'est 10 révélé faiblement immunogène chez les enfants de moins de deux ans, alors que le polysaccharide de Nm groupe B est non immunogène chez l'homme et partage des épitopes avec des glycoprotéines d'adhésion présentes dans les cellules neuronales humaines.

15 Il n'existe donc pas de vaccin universel capable de prévenir les infections provoquées par l'ensemble des sérogroupes des méningocoques et capable de répondre à la variabilité antigénique propre aux pathogènes bactériens en général et à Nm en particulier.

20 En raison de la réactivité croisée du polysaccharide groupe B de Nm avec les antigènes humain, de la multiplicité des sérogroupes et de la variabilité antigénique de Nm, les stratégies proposées à ce jour ne peuvent conduire à un vaccin efficace dans toutes les 25 situations.

Les recherches se sont alors concentrées sur l'étude d'éléments caractéristiques responsables de la spécificité de la pathogénèse méningococcique.

30 La plupart des gènes qui ont été étudiés dans l'une quelconque des deux bactéries Nm ou Ng possèdent leur homologue dans la deuxième bactérie.

De la même manière, la plupart des facteurs de virulence jusqu'ici identifiés dans Nm ont une contrepartie dans Ng, c'est-à-dire la piline, les

protéines PilC, les protéines d'opacité et les récepteurs de la lactoferrine et de la transferrine.

Les attributs spécifiques des méningocoques caractérisés dans l'art antérieur sont la capsule, les protéines Frp analogues aux toxines RTX, les protéines de la membre externe Opc, la peroxydase glutathione, la porine PorA et le gène rotamase.

10 Parmi ceux-ci, seule la capsule est invariablement présente dans les souches virulentes de Nm. Cependant, de nombreux pathogènes extra-cellulaires possèdent une capsule sans pour autant traverser la barrière hémato-encéphalique.

15 Des attributs non encore identifiés doivent donc être responsables de la spécificité de la pathogénèse méningococcale. Ces attributs sont vraisemblablement codés par des séquences d'ADN présentes parmi les méningocoques mais absentes chez les gonocoques.

20 Les inventeurs ont développé une nouvelle voie d'approche basée sur l'isolement soustractif des gènes Nm-spécifiques, ces gènes devant être liés à la pathogénèse spécifique de Nm, et, plus particulièrement au franchissement de la barrière hémato-encéphalique.

25 La méthode soustractive développée dans l'art antérieur a abouti à la production de marqueurs épidémiologiques pour certains isolats de Nm. Ces marqueurs sont d'une utilité limitée : ils ne couvrent pas l'ensemble des sérogroupes de l'espèce Nm.

30 Par contraste avec ces études, les travaux des inventeurs ont conduit, en confrontant Nm à l'ensemble du chromosome de Ng, cisaillé de manière aléatoire, à la mise au point de moyens pour cloner l'ensemble des ADN présents chez Nm et absents chez Ng, fournissant ainsi des outils de haute spécificité vis-à-vis de Nm et permettant ainsi de répondre pour la première fois à la 35 variabilité génétique de l'espèce.

Les termes "présent" et "absent", tel qu'utilisés dans la description et les revendications en rapport avec les ADN d'une souche, ou leurs produits d'expression, sont appréciés par rapport à des conditions d'hybridation identiques (16h à 65°C, avec NaPO₄ 0,5M, pH 7,2; EDTA-Na 0,001M, 1%, 1% d'albumine de sérum bovin et 7% de dodécylsulfate de sodium), en utilisant une même sonde et une même intensité de marquage de la sonde, une même quantité d'ADN chromosomique et un même élément de comparaison (ADN chromosomique de la souche homologue).
Ainsi, on considère que l'ADN est présent lorsque le signal obtenu avec la sonde est pratiquement le même que celui obtenu avec la souche de référence.

En revanche, on considère que l'ADN est absent lorsque ce signal apparaît très faible.

Une deuxième considération sur les pathogénicités de Nm et de Ng conduit à prendre en compte leur aptitude commune à coloniser et à pénétrer la muqueuse puis à envahir l'espace sous-épithérial de cette dernière. Il est fort vraisemblable que ce processus implique des facteurs de virulence communs aux deux pathogènes. A cet égard, on sait qu'un certain nombre de facteurs de virulence ont été déjà identifiés chez Nm et chez Ng, comme les protéines pili, PilC, les protéines d'opacité, les protéases d'IgA, les protéines de liaison à la transferrine et à la lactoferrine, et des lipooligosaccharides.

La démarche des inventeurs s'est donc étendue à la recherche de régions de Nm, spécifiques de Nm et de Ng, mais absentes chez l'espèce non pathogène Nl, et d'une manière générale à la recherche, par les moyens mis au point conformément à l'invention, de régions chromosomiques d'ADN et de leurs produits d'expression, spécifiques d'une espèce donnée.

L'invention a donc pour but de fournir des ADN de Nm spécifiques de son pouvoir pathogène et des moyens pour les obtenir, notamment en élaborant des banques formées exclusivement de ces ADN Nm-spécifiques.

5 Elle vise également les produits dérivés de ces séquences d'ADN.

L'invention vise également la mise à profit des caractères spécifique et exhaustif de ces banques pour élaborer des outils utilisables notamment en diagnostic, 10 thérapie et prévention.

Les ADN de l'invention sont caractérisés en ce qu'il s'agit de tout ou partie de gènes, avec leur phase de lecture, présents chez *Neisseria meningitidis*, mais absents soit chez *Neisseria gonorrhoeae*, soit chez 15 *Neisseria lactamica*, à l'exception des gènes impliqués dans la biosynthèse de la capsule polysaccharidique, *frpA*, *frpC*, *opc*, *por A*, rotamase, de la séquence IS1106, des IgA protéases, de la piline, de *pilC*, des protéines qui lient la transferrine et des protéines d'opacité.

20 Comme précisé plus haut, les termes "présents" et "absents" sont appréciés par rapport aux conditions d'hybridation telles qu'utilisées dans les Southern blots décrits dans les exemples et rappelées plus haut.

25 On notera que ces ADN englobent les variants dès lors qu'ils expriment une fonction propre à l'espèce Nm, plus particulièrement un phénotype retrouvé uniquement chez Nm ou en commun exclusivement avec Ng.

30 Selon un aspect majeur, ces ADN sont spécifiques de la pathogénicité de *Neisseria meningitidis* et ce, en dépit de la variabilité génétique de cette espèce.

Selon un mode de réalisation de l'invention, lesdits ADN sont spécifiques de Nm par rapport à Ng.

Plus particulièrement, les ADN Nm-spécifiques sont absents de *Neisseria lactamica* et de *Neisseria cinerea*.

De façon surprenante, la majorité des différences génétiques entre les souches de méningocoques et celles de gonocoques apparaissent regroupées en régions distinctes, qui correspondraient à des îlots de 5 pathogénicités comme précédemment décrit pour *E. coli* et *Y. pestis*.

Ainsi, dans une disposition préférée de l'invention, ces ADN sont également caractérisés en ce qu'ils comprennent une ou plusieurs séquence(s), telle(s) que 10 présente(s) sur le chromosome de *Neisseria meningitidis* Z2491 entre *tufA* et *pilT*, ou région 1 du chromosome, et/ou la ou les séquence(s) capable(s) de s'hybrider avec la ou les séquence(s) ci-dessus, sous réserve d'être spécifique(s) de *Neisseria meningitidis*.

15 Par "spécifique", on désigne dans la description et les revendications les séquences de nucléotides qui ne s'hybrident qu'avec celles de Nm, dans des conditions d'hybridation données dans les exemples et rappelées plus haut.

20 On notera à cet égard que, de manière générale, lorsqu'on fait référence dans la description et les revendications à "tout ou partie" d'une séquence, cette expression doit être appréciée par rapport à la spécificité définie ci-dessus.

25 De même, tout ou partie d'un peptide, ou un fragment d'un peptide ou d'un anticorps désigne un produit présentant les propriétés biologiques respectivement du peptide natif ou de l'anticorps formé contre le peptide.

30 Dans la région 1, sont regroupés des gènes de la capsule de *Neisseria meningitidis*.

Des ADN de ce type présentent une séquence correspondant, pour tout ou partie, à SEQ ID N°9, 13, 22 ou 30, et/ou à toute séquence se situant à plus ou moins 20 kb de ces SEQ ID sur le chromosome d'une souche de Nm, 35 et/ou présentent une séquence capable de s'hybrider avec

au moins un fragment de l'une quelconque de ces séquences.

Dans une autre disposition préférée de l'invention, ces ADN sont également caractérisés en ce qu'ils sont 5 constitués par une ou plusieurs séquence(s), telle(s) que présente(s) sur le chromosome de *Neisseria meningitidis* Z2491 entre *pilQ* et *λ740*, ou région 2 du chromosome, et/ou la ou les séquence(s) capable(s) de s'hybrider avec la ou les séquence(s) ci-dessus, sous réserve d'être 10 spécifique(s) de *Neisseria meningitidis*.

Des ADN selon cette disposition présentent une séquence correspondant, pour tout ou partie, à SEQ ID N°1, 2, 4, 6, 7, 10, 15, 31 ou 34, et/ou à toute séquence 15 se situant à plus ou moins 20 kb de ces SEQ ID sur le chromosome d'une souche de Nm, et/ou présentent une séquence capable de s'hybrider avec au moins un fragment de l'une quelconque de ces séquences.

L'invention vise tout spécialement tout ou partie de la 20 séquence d'ADN SEQ ID N°36 de 15620 pb, et les séquences correspondant aux cadres ouverts de lecture SEQ ID N°37, SEQ ID N° 38, SEQ ID N° 39, SEQ ID N° 40, SEQ ID N° 41, SEQ ID N° 42, SEQ ID N° 43, SEQ ID N° 44 et SEQ ID N° 45.

Dans encore une autre disposition préférée de 25 l'invention, ces ADN sont également caractérisés en ce qu'ils sont constitués par une ou plusieurs séquence(s), telle(s) que présente(s) sur le chromosome de *Neisseria meningitidis* Z2491 entre *argF* et *opaB*, ou région 3 du chromosome, et/ou la ou les séquence(s) capable(s) de 30 s'hybrider avec la ou les séquence(s) ci-dessus, sous réserve d'être spécifique(s) de *Neisseria meningitidis*.

Des ADN selon cette disposition sont caractérisés en ce qu'ils présentent une séquence correspondant pour tout ou partie à SEQ ID N°8, 21, 23, 25, 26, 28, 29, 32 ou 35, 35 et/ou à toute séquence se situant à plus ou moins 20 kb

de ces SEQ ID sur le chromosome d'une souche de Nm, et/ou, présentent une séquence capable de s'hybrider avec au moins un fragment de l'une quelconque de ces séquences.

5 Les régions 1, 2, 3, identifiées ci-dessus, présentent une forte proportion de séquences *Neisseria meningitidis* spécifiques, et entrent également dans le cadre de l'invention.

10 D'autres ADN représentatifs de la spécificité vis-à-vis de *Neisseria meningitidis* présentent une ou plusieurs séquences telle(s) que présente(s) sur le chromosome de *Neisseria meningitidis* Z2491, mais ne font pas partie des régions 1, 2, 3 définies ci-dessus.

15 De tels ADN comprennent une ou plusieurs séquences correspondant pour tout ou partie à SEQ ID n°3, 5, 11, 12, 14, 16, 18, 19, 20, 24, 27 ou 33, et/ou à toute séquence se situant à plus ou moins 20 kb de ces SEQ ID sur le chromosome d'une souche de Nm, et/ou présentent une séquence capable de s'hybrider avec de telles 20 séquences.

Compte tenu des applications particulièrement visées, l'invention concerne plus spécialement les ADN ci-dessus impliqués dans la pathogénèse de l'organisme bactérien.

25 Elle vise, en particulier, les ADN répondant à au moins l'une des caractérisations données ci-dessus, et codant pour une protéine exportée au-delà de la membrane cytoplasmique et/ou dont tout ou partie de leur séquence correspond à la région conservée desdits ADN.

30 Ainsi, selon un autre mode de réalisation de l'invention, les ADN sont communs avec ceux de Ng, mais sont absents de chez Nl.

Il s'agit plus spécialement d'ADN présents sur la région 4 (arg J à reg F) ou sur la région 5 (marqueur 35 lambda 375 à pen A) sur le chromosome de Nm Z2491 et/ou

capables de s'hybrider avec lesdits ADN présents, sous réserve d'être spécifiques de Nm et de Ng par rapport à N1.

5 Par "spécifique de Nm et de Ng par rapport à N1", on désigne des ADN qui s'hybrident avec les ADN de Nm et de Ng dans les conditions d'hybridation des exemples (voir en particulier l'exemple 4).

10 Les ADN des régions 4 et 5, et ceux capables de s'hybrider avec ces ADN, sous réserve d'exprimer les fonctions propres à Nm, présentent l'avantage d'intervenir de manière majeure dans la virulence de Nm, en étant impliqués dans l'étape de colonisation et de pénétration initiales et dans la dissémination septicémique.

15 Selon d'autres dispositions, l'invention vise les vecteurs de transfert et d'expression, tels que plasmides, cosmides ou bactériophages, comportant au moins un ADN tel que défini ci-dessus.

20 Elle vise aussi les cellules hôtes telles que transformées par au moins un ADN tel que défini ci-dessus.

25 D'autres cellules hôtes de l'invention comportent des gènes ou des fragments de gènes spécifiques de Nm et sont caractérisées en ce que leur chromosome est déléte d'au moins un ADN selon l'invention, en particulier d'un ADN responsable de la pathogénicité. Il s'agit plus spécialement de cellules bactériennes, notamment de Nm.

30 L'invention a également pour objet les ARN dont la séquence correspond pour tout ou partie à la transcription d'au moins une séquence ou fragment de séquence d'ADN tel que défini ci-dessus.

Les acides nucléiques anti-sens des ADN tels que définis ci-dessus, ou de fragments de ces ADN, font également partie de l'invention.

Ces acides nucléiques anti-sens portent le cas échéant au moins un substituant telle qu'un groupe méthyle et/ou un groupe glycosyle.

5 D'autres produits entrant dans le champ de l'invention sont constitués par des polypeptides.

Ces polypeptides sont caractérisés en ce qu'ils présentent un enchaînement d'acides aminés correspondant à tout ou partie d'une séquence telle que codée par les 10 acides nucléiques définis dans ce qui précède, ou telle que déduite des séquences de ces acides nucléiques.

Il s'agit avantageusement de polypeptides correspondant à tout ou partie de polypeptides exportés au-delà de la membrane cytoplasmique, plus spécialement de polypeptides correspondant à tout ou partie de ceux 15 tels que codés par une région conservée.

En variante, les polypeptides de l'invention peuvent être modifiés par rapport à ceux correspondant aux 20 séquences d'acides nucléiques, et ce de manière à être particulièrement adaptés pour une application donnée, en particulier une application vaccinale.

Par modification, on entend toute altération, déletion, substitution chimique, dès lors qu'elle 25 n'affecte pas les propriétés biochimiques des polypeptides natifs correspondants, plus spécialement des protéines fonctionnelles telles qu'exportées au niveau du périplasme et de la membrane externe.

D'autres produits conformes à l'invention sont constitués par les anticorps dirigés contre les polypeptides ci-dessus.

30 L'invention vise ainsi les anticorps polyclonaux, ainsi que les anticorps monoclonaux, caractérisés en ce qu'ils reconnaissent au moins un épitope d'un polypeptide tel qu'évoqué plus haut.

Elle vise également les fragments de ces anticorps, 35 plus particulièrement les fragments Fv, Fab, Fab'2.

Les anti-anticorps capables de reconnaître les anticorps définis ci-dessus, ou leurs fragments, selon une réaction de type antigène-anticorps, font également partie de l'invention.

5 Conformément à l'invention, les différents produits considérés ci-dessus sont obtenus par voie de synthèse et/ou biologique en opérant selon les techniques classiques.

10 Les acides nucléiques peuvent être également obtenus à partir de banques constituées d'ADN Nm- spécifiques, telles qu'élaborées selon une technique soustractive, cette technique comprenant :

15 - le mélange de deux populations d'ADN,
- la réalisation d'au moins une itération d'hybridation-amplification soustractive, et
- la récupération du ou des ADN souhaités, suivie le cas échéant de leur purification avec l'élimination des séquences redondantes.

20 Conformément à l'invention, les deux populations d'ADN proviennent respectivement d'une souche de *Neisseria meningitidis*, dite souche de référence, pour laquelle la banque spécifique doit être constituée, et d'une souche de *Neisseria*, dite souche de soustraction, présentant une homologie en séquences primaires d'ADN 25 supérieure à environ 70% avec la souche de *Neisseria meningitidis*, les séquences d'ADN des souches de soustraction et de référence étant telles qu'obtenues respectivement par cisaillement aléatoire, et par clivage par une endonucléase de restriction capable de produire 30 des fragments de taille inférieure à environ 1kb.

L'invention vise en particulier un procédé d'obtention de banques d'ADN *Neisseria meningitidis* spécifiques, comportant les étapes de :

35 - cisaillement aléatoire de l'ADN chromosomique d'une souche *Neisseria gonorrhoeae*, dite souche de

soustraction, notamment par passages répétés à travers une seringue,

- clivage de l'ADN chromosomique d'une souche de *Neisseria meningitidis*, dite souche de référence, de préférence par une enzyme de restriction produisant des fragments de taille inférieure à 1kb environ,

- ligature des fragments d'ADN de la souche de référence, clivés par l'enzyme de restriction, avec des amorces oligonucléotidiques appropriées,

10 - réalisation d'une itération d'hybridation-amplification soustractive par :

. mélange des deux populations d'ADN dans des conditions appropriées pour l'hybridation des séquences homologues, puis

15 . amplification des fragments auto-réannelés et récupération de ces fragments,

. digestion de ces fragments par une enzyme de restriction, et re-ligature à des amorces oligonucléotides suivie d'une

20 - purification de l'ADN ligaturé, et le cas échéant, d'une nouvelle itération d'hybridation soustractive, comportant le mélange de fragments d'ADN de *Neisseria gonorrhoeae* cisailé comme indiqué ci-dessus avec les fragments d'ADN de *Neisseria meningitidis* issus 25 de l'itération précédente, suivi, si on le souhaite du clonage des ADN de la banque.

Les amorces utilisées sont des amorces oligodésoxynucléotidiques adaptées à l'endonucléase de restriction utilisée et permettant une insertion dans un 30 site de clonage, tel que le site EcoRI du plasmide pBluescript. On choisira avantageusement de telles amorces parmi les oligodésoxynucléotides référencés dans le listing de séquence sous SEQ ID n°36 à 45.

Les banques ainsi obtenues sont formées d'ADN spécifiques des méningocoques et absents chez les gonocoques.

La spécificité des ADN a été vérifiée comme exposé dans les exemples, à chaque itération par Southern blots, avec des gènes communs à la souche de soustraction et à la souche de référence, ou avec l'ADN total de chacune des souches digéré par une endonucléase de restriction, telle que *Cla*I.

10 A chaque itération, a également été vérifiée l'exhaustivité de la banque d'ADN par Southern blotting avec des sondes connues pour être spécifiques de la souche de référence, à savoir pour *Neisseria meningitidis*, les gènes *frp*, *opc*, rotamase, notamment.

15 Les expériences réalisées ont montré que les banques obtenues selon le procédé de l'invention sont dépourvues des gènes présentant une homologie significative avec des espèces de *Neisseria* autre que *Neisseria meningitidis*, par exemple les gènes, *ppk* ou *pilC1*, et ce généralement, 20 en seulement 2 ou 3 itérations.

Si nécessaire, deux voies, non exclusives l'une de l'autre, peuvent être empruntées.

25 Il est possible de procéder à une $(n+1)^{\text{ème}}$ itération, en utilisant l'ADN de l'itération n comme population d'ADN de la souche de référence.

En variante, on réalise une deuxième banque, indépendante de la première, avec une enzyme de restriction de spécificité différente de celle utilisée dans la première banque, par exemple *Mbo*I.

30 Dans tous les cas, il est préférable de conserver chacun des produits issus de chacune des itérations réalisées.

L'invention vise également l'utilisation de la technique soustractive décrite ci-dessus pour obtenir des

banques d'ADN communs entre Nm et Ng, mais spécifiques par rapport à Nl.

On constitue avantageusement trois banques différentes, dont deux par digestion de l'ADN chromosomique de Nm par *MboI* et *Tsp5091*, et la troisième, par digestion de l'ADN chromosomique de Nm avec *MspI*. Deux séries de soustraction permettent de récupérer des ADN présentant la spécificité recherchée, comme décrit dans les exemples.

Le procédé d'obtention de ces banques et les banques elles-mêmes font également partie de l'invention.

On observera que, de manière générale, le procédé de l'invention est applicable pour l'obtention de banques d'ADN spécifiques d'une espèce de cellule donnée ou d'un variant donné d'une même espèce, dès lors qu'il existe une autre espèce ou un autre variant proche génomiquement et exprimant des pouvoirs pathogènes différents.

En appliquant le procédé de l'invention, on constituera avantageusement des banques d'ADN spécifiques d'espèces données de cryptocoques, d'*Haemophilus*, de pneumocoques ou encore d'*Escherichia coli*, ou plus généralement de tout agent bactérien appartenant à la même espèce et disposant de pathovars différents.

De même, à partir de ces banques, l'invention fournit les moyens de disposer de facteurs de virulence spécifiques d'une espèce ou d'un variant donné.

De telles banques constituent donc des outils présentant un intérêt majeur pour disposer d'attributs responsables de la spécificité d'un pathogène, cette application étant plus spécialement illustrée conformément à l'invention par l'obtention de banques renfermant les attributs responsables de la spécificité de la pathogénèse méningococcique.

L'étude des produits de l'invention, acides nucléiques, polypeptides et anticorps, a permis de mettre

en évidence une spécificité absolue vis-à-vis de *Neisseria meningitidis*, quelle que soit la souche et sa variabilité.

5 Ces produits sont donc particulièrement appropriés pour le diagnostic ou la prévention des infections et méningites provoquées par *Neisseria meningitidis*, que ce soit chez l'adulte ou l'enfant et quel que soit le sérogroupe de la souche en cause.

10 La méthode de diagnostic, selon l'invention, d'une infection méningococcique, et plus particulièrement de la méningite méningococcique, par mise en évidence de la présence de *Neisseria meningitidis* dans un échantillon à analyser, est caractérisé par les étapes de :

- 15 - mise en contact, d'un échantillon à analyser, à savoir un échantillon biologique ou une culture cellulaire, avec un réactif élaboré à partir d'au moins un acide nucléique tel que défini ci-dessus, le cas échéant sous forme de sonde nucléotidique ou d'amorce, ou en variante à partir d'au moins un anticorps, ou un fragment d'anticorps, tel que défini ci-dessus, dans des conditions permettant respectivement une hybridation ou une réaction de type antigène-anticorps, et
- 20 - révélation du produit de réaction éventuellement formé.

25 Lorsque le réactif est élaboré à partir d'un acide nucléique, celui-ci peut se présenter sous forme de sonde nucléotidique dans laquelle l'acide nucléique, ou un fragment de ce dernier, est marqué afin de permettre sa révélation. Des marqueurs appropriés comprennent des 30 marqueurs radio-actifs, fluorescents, enzymatiques ou luminescents.

En variante, l'acide nucléique est inclus dans une cellule hôte, utilisée comme réactif.

Dans ces différentes formes, l'acide nucléique est utilisé tel quel ou sous forme d'une composition avec des véhicules inertes.

5 Lorsque le réactif est élaboré à partir d'un anticorps, ou d'un fragment d'anticorps, celui-ci peut être marqué aux fins de révélation. Le plus couramment, on utilise un marqueur fluorescent, enzymatique, radio-actif ou luminescent.

10 L'anticorps, ou le fragment d'anticorps utilisé, le cas échéant, marqué, peut être utilisé tel quel ou sous forme d'une composition avec des véhicules inertes.

L'échantillon utilisé dans l'étape de mise en contact est un échantillon biologique, issu d'un mammifère, tel que liquide céphalo-rachidien, urine, sang, salive.

15 L'étape de révélation est réalisée dans des conditions permettant de mettre en évidence le produit de réaction lorsqu'il s'est formé. Des moyens classiques mettent en oeuvre des réactions de fluorescence, luminescence, colorées, radio-actives ou encore des techniques d'autoriadographie. Il est également possible de quantifier le produit.

20 Les produits marqués, acides nucléiques et anticorps font également partie en tant que produits nouveaux de l'invention.

La méthode définie ci-dessus peut être appliquée au diagnostic d'une réaction immunitaire spécifique d'une infection méningococcique.

25 On utilise alors comme réactif un polypeptide conforme à l'invention, tel que codé par lesdites séquences d'acides nucléiques, correspondant au produit natif, ou un polypeptide modifié, mais possédant l'activité biologique et immunologique de polypeptide natif correspondant.

Il s'agit avantageusement d'un polypeptide tel qu'exporté au-delà de la membrane cytoplasmique de *Neisseria meningitidis*, plus particulièrement de la partie d'un tel polypeptide correspondant à la région 5 conservée de l'ADN.

L'invention vise également des kits pour la mise en œuvre des méthodes définies ci-dessus. Ces kits sont caractérisés en ce qu'ils comportent :

- au moins un réactif tel que défini ci-dessus, à savoir de type acide nucléique, anticorps ou polypeptide,
- les produits, notamment marqueurs ou tampons, permettant la réalisation de la réaction d'hybridation nucléotidique ou de la réaction immunologique visée, ainsi qu'une notice d'utilisation.

La spécificité des produits de l'invention et leur localisation sur le chromosome de *Neisseria meningitidis* Z2491 soit regroupés en région, pouvant être interprétées comme des îlots de pathogénicité, soit isolés sur le chromosome, leur confèrent un intérêt tout particulier pour la réalisation de compositions vaccinales à visée universelle, c'est-à-dire quelque soit la souche et la variabilité qu'elle exprime. Ces compositions peuvent inclure dans leur spectre d'autres prophylaxies, et être, par exemple, associées aux vaccins de l'enfance.

L'invention vise donc des compositions vaccinales incluant dans leur spectre une prophylaxie à visée antiméningococcique, destinées à prévenir toute infection susceptible d'être provoquée par *Neisseria meningitidis*, ces compositions étant caractérisées en ce qu'elles comprennent, en association avec un ou des véhicule(s) physiologiquement acceptable(s), une quantité efficace de polypeptides ou d'anti-anticorps ou de leurs fragments tels que définis ci-dessus, ces produits étant éventuellement conjugués, afin de renforcer leur 35 immogénicité.

Des molécules immunogènes utilisables comprennent la protéine de polyovirus, la toxine tétanique, ou encore la protéine issue de la région hypervariable d'une piline.

En variante, les compositions vaccinales selon 5 l'invention sont caractérisées en ce qu'elles comprennent, en association avec un/des véhicule(s) physiologiquement acceptable(s), une quantité efficace :

- d'acides nucléiques tels que définis ci-dessus,
- de cellules hôtes transformées telles que définies 10 plus haut, ou

- de cellules de Nm dont le chromosome a été déléte 15 d'au moins une séquence d'ADN selon l'invention impliquée dans la pathogénicité de la bactérie. Le matériel nucléotidique utilisé est avantageusement placé sous le contrôle d'un promoteur favorisant son expression in vivo et la synthèse de la protéine correspondante. Il est également possible afin de renforcer l'immunogénicité, d'associer ce matériel nucléique avec un ADN ou un ARN encodant une molécule porteuse telle que protéine de 20 polyovirus, toxine tétanique, protéine issue de la région hypervariable d'une piline.

Les compositions vaccinales de l'invention sont administrables par voie parentérale, sous-cutanée, intramusculaire ou encore sous forme de spray.

25 D'autres caractéristiques et avantages de l'invention sont donnés dans les exemples qui suivent afin d'illustrer celle-ci sans toutefois en limiter sa portée.

Dans ces exemples, il sera fait référence aux 30 figures 1 à 11 qui représentent respectivement

- les figures 1A, 1B, 1C, 1D, 1E, 1F et 1G l'analyse de la banque soustractive Tsp5091,
- la figure 2, la distribution de séquences Nm-spécifiques par rapport à Ng sur le chromosome de la

souche Z2491, (partie gauche) et de séquences Nm spécifiques par rapport à Nl (partie droite),

- la figure 3A à 3C, la réactivité des clones des 3 régions du chromosome, selon l'invention, envers une panel de souches du genre *Neisseria*.

- la figure 4, la position, dans la région 2 du chromosome de Nm, d'oligonucléotides utilisés comme sondes,

- les figures 5, 6 et 7, les Southern blots d'un panel de souches du genre *Neisseria*, en utilisant des parties de la région 2 de Nm comme sondes,

- les figures 8A à 8C, les Southern blots avec 3 banques soustractive sur un panel de 12 souches de *Neisseria*, et

- les figures 9, 10 et 11, la réactivité de clones des 3 banques soustractive vis-à-vis de Nm, Nl et Ng.

Dans les exemples qui vont suivre, les matériels et méthodes suivants ont été utilisés :

Souches bactériennes - Pour la réalisation des banques soustractive, on a utilisé la souche Z2491 de Nm (Achtman et al., 1991, *J. Infect. Dis.* 164, 375-382) les souches MS11 (Swanson et al., 1974, *Infect. Immun.* 10, 633-644), et les souches 8064 et 9764 de Nl, étant entendu que tout autre souche de l'espèce considérée pourrait être utilisée.

Afin de vérifier la spécificité de ces banques, 6 souches de Nm, 4 souches de Ng, une souche de Nl (*Neisseria lactamica*) et une souche de Nc (*Neisseria cinerea*) ont été utilisées.

Les six souches de Nm sont : Nm Z2491 de sérogroupe A, Nm 8013 de sérogroupe C (XN collection), Nm 1121 non sérogrouppable (XN collection), Nm 1912 sérogroupe A (XN collection), Nm 7972 de sérogroupe A (XN collection) et Nm 8216 de sérogroupe B (XN collection).

Les quatre souches de Ng sont : Ng MS11 (Institut Pasteur, Paris), Ng 403 (Institut Pasteur, Paris), Ng

6934 (Institut Pasteur, Paris), Ng WI (isolée à partir d'une infection gonococcique disséminée), Ng 4C1, Ng 6493 et Ng FA 1090.

5 Les souches de N1 sont N1 8064 et N1 9764 (XN collection) et celle de Nc, Nc 32165 (XN collection).

Techniques de génétique moléculaire

Sauf indication contraire, les techniques et réactifs utilisés correspondent à ceux recommandés par Sambrook et al (Sambrook et al 1989, Molecular Cloning: 10 A Laboratory Manual. Cold Spring Harbor Laboratory Press). Les oligodésoxynucléotides utilisés dans cette étude sont :

RBam12, 3'AGTGGCTCCTAG 54 (SEQ ID N°54)

15 RBam24, 5' AGCACTCTCCAGCCTCTCACCGAG 3'; (SEQ ID N°55)

Jbam12, 3' GATCCGTTCATG 5'; (SEQ ID N°60)

JBAM24, 5' ACCGACGTCGACTATCCATGAACG 3'; (SEQ ID N°61)

REco12, AGTGGCTCTTAA; (SEQ ID N°56)

REco24, 5' AGCACTCTCCAGCCTCTCACCGAG 3'; (= RBam 24)

20 JEco12, GTACTTGCTTAA; (SEQ ID N°62)

JEco24, 5' ACCGACGTCGACTATCCATGAACG 3'; (= JBam24)

NEco12, AATTCTCCCTCG; (SEQ ID N°64)

NEco24, AGGCAACTGTGCTATCCGAGGGAG; (SEQ ID N°65).

25 Transferts sur membranes (Southern blots)

Les transferts sur membranes ont été réalisés par transferts capillaires sur des membranes en nylon chargées positivement (Boehringer Mannheim). Les hybridations ont été réalisées à 65°C dans une solution 30 comprenant NaPi 0,5M pH7,2/EDTA 1mM/SDS 7%/ BSA 1%. Les lavages des membranes ont été réalisées dans une solution comprenant NaPi 40mM pH7,2/EDTA 1mM/SDS 1%. Le lavage final a été réalisé à 65°C pendant 5 min.

La sonde frp, obtenue avec des oligonucléotides 35 basés sur la séquence de frpA correspond à 2,4 kb de

l'extrémité 5' du gène de la souche Z2491. Les sondes *opc* et *rotamase* correspondant aux gènes entiers sont produites à partir de la souche Z2491 en utilisant des oligonucléotides réalisés sur la base de séquences publiées. Les sondes *pilC1* et *ppk* (polyphosphate kinase) correspondent aux inserts des plasmides pJL1 et pBluePPK6001, respectivement.

10 Exemple 1 : Réalisation de banques d'ADN présents chez Nm et absents chez Ng.

a. Banque "MboI"

15 Réalisation - L'ADN de Nm Z2491 a été clivé par l'endonucléase *MboI* et soumis à deux itérations d'une méthode, appelée ci-après CDA (Comprehensive Difference Analysis). Cette méthode comprend une hybridation soustractive en présence d'un excès d'ADN cisaillé de Ng MS11 et une amplification par PCR de celles des séquences méningococciques qui, étant absentes de ou ne présentant 20 pas d'homologie significative avec l'ADN de Ng MS11, pouvaient se ré-anneeler.

25 L'ADN chromosomique de la souche Ng MS11 est cisaillé de manière aléatoire par passages répétés à travers une seringue hypodermique jusqu'à obtention de fragments dont la taille s'échelonne de 3 à 10 kb. Ces fragments d'ADN sont purifiés par extraction phénolique.

30 L'ADN chromosomique de la souche Nm Z2491 est, quant à lui, clivé par l'endonucléase de restriction *MboI*. Ces fragments d'ADN (20 µg) sont ligaturés à 10 nmoles des oligonucléotides annelés RBam12 et RBam24. Les amorces en excès sont éliminées par électrophorèse sur un gel d'agarose à 2% à bas point de fusion. La partie du gel 35 contenant des fragments amplifiés de taille supérieure à 200 pb est excisée et digérée par la β -agarase. Ces fragments sont purifiés par extraction phénolique.

Afin de réaliser une hybridation soustractive (première itération), 0,2 µg d'ADN Nm, ligaturé aux oligonucléotides RBam, est mélangé à 40 µg d'ADN Ng dans un volume total de 8 ml d'un tampon EE 3X (un tampon EE 5 IX est composé de N-(2-hydroxyéthyl) pipérazine-N'-(acide sulphonique propane 3) 10 mM et d'EDTA 1 mM, son pH est de 8.0). Cette solution est recouverte d'huile minérale et l'ADN est dénaturé par chauffage à 100°C pendant 2 min. 2 µl de NaCl 5M sont ajoutés et on laisse le mélange 10 s'hybrider à 55°C pendant 48h. Le mélange réactionnel est dilué à 1/10 dans une solution préchauffée composée de NaCl et de tampon EE, puis immédiatement placé sur de la glace.

10 µl de cette dilution sont ajoutés à 400 µl de 15 mélange réactionnel pour PCR (Tris.HCl pH9.0 10mM; KCl 50 mM; MgCl₂ 1,5 mM; Triton X100 0,1 %; 0,25 mM de chacun des quatre désoxynucléotides triphosphate ; Taq 20 polymérase 50 unités par ml). Le mélange est incubé pendant 3 min à 70°C pour compléter les extrémités des fragments ré-annelés d'ADN méningococciques.

Après dénaturation à 94°C pendant 5 min et addition de l'oligonucléotide RBam24 à raison de 0,1 nmole par 100 µl, les hybrides sont amplifiés par PCR (30 cycles de 1 min à 94°C, 1 min à 70°C et 3 min à 72°C suivis par 25 1 min à 94°C et 10 min à 72°C; Perkin-Elmer GeneAmp 9600).

Les fragments méningococciques amplifiés sont 30 séparés sur gel des amorces et des ADN gonococciques de hauts poids moléculaires. Ils sont digérés par MboI et de nouveaux oligonucléotides JBam12 et JBam24 leur sont 35 ligaturés. Ces ADN ligaturés sont à nouveau purifiés sur gel et extraits au phénol.

Une seconde itération d'hybridation soustractive est réalisée sur 40 µg d'ADN Ng cisaillé de manière aléatoire et 25 ng d'ADN ligaturé aux oligonucléotides JBam tel

qu'obtenu à l'issue de la première itération d'hybridation soustractive. Lors de cette seconde itération, l'amplification de l'ADN Nm auto-annelé est réalisée à l'aide de l'oligonucléotide Jbam24.

5 Spécificité - Afin de confirmer leur Nm-spécificité, les séquences amplifiées après la seconde itération de la méthode CDA sont marquées et utilisées comme sonde pour de l'ADN digéré par *Cla*I issu d'un panel de six souches de *Neisseria meningitidis*, quatre de 10 *Neisseria gonorrhoeae*, une de *Neisseria lactamica* et une de *Neisseria cinerea*.

15 Les Southern blots réalisés montrent que les séquences amplifiées à l'issue de la seconde itération de la méthode CDA présentent une forte réactivité avec de nombreuses bandes correspondant aux meningocoques et ne 20 présentent pas de réactivité avec les bandes correspondant aux souches Ng, Nl, Nc.

La banque "MboI" apparaît donc comme Nm-spécifique.

25 Exhaustivité - Afin de tester l'exhaustivité de la banque, l'ensemble des produits issus de la première et de la seconde itérations de la méthode CDA ainsi que les matériaux chromosomiques initiaux de Nm Z2481 et de Ng MS11 sont soumis à électrophorèse sur gel d'agarose, transférés sur membrane et mis en contact avec des sondes 30 comprenant des gènes connus pour être méningococcus-spécifiques, à savoir *frp*, *opc*, *rotamase* (Southern blot).

Il résulte de ces hybridations que le gène Nm-spécifique *frp* est représenté dans la banque MboI par un fragment de 600 pb, mais qu'aucune activité n'est 35 observée pour les gènes *rotamase* et *opc*. La banque MboI, bien que Nm-spécifique, ne peut donc être considérée comme exhaustive.

Etant donné leur haute spécificité, les fragments issus de la seconde itération de la méthode CDA pour la

banque *MboI* peuvent néanmoins être clonés sur le site *BamHI* du plasmide *pBluescript*.

Une séquence correspondant à un quelconque des gènes *Nm*-spécifiques ne peut être incluse dans la banque soustractive que si elle est portée par un fragment de restriction de taille appropriée. Cette condition est fonction de deux facteurs. Premièrement, la probabilité pour que les plus grands fragments soient entièrement *Nm*-spécifiques est faible. Deuxièmement, même si de tels fragments existaient, ils seraient sous-représentés dans la banque du fait des limitations de la technique PCR dont l'efficacité d'amplification diminue avec l'augmentation de la taille des fragments. Les fragments de taille supérieure à environ 600 pb ne sont pas inclus dans la banque. Du fait de l'absence, dans le chromosome de *Nm* Z2491, de fragments *Mbo* de taille appropriée, les gènes *rotamase* et *opc* ne peuvent être inclus dans la banque. Une enzyme quelconque ne peut à elle seule produire un petit fragment correspondant à un gène *Nm*-spécifique quelconque. Une deuxième banque a donc été réalisée en utilisant une autre enzyme de restriction avec une spécificité différente : *Tsp509*.

b. Banque "Tsp509I"

25 Réalisation - L'enzyme *Tsp509I* présente l'avantage de produire des fragments de plus petite taille (inférieure à 1 kb environ) que l'enzyme *MboI*.

30 *Tsp509I* reconnaît la séquence AATT et laisse, en saillie en 5', une séquence de 4 bases compatible avec *EcoRI*. Les oligonucléotides utilisés sont *Reco*, *Jeco* et *NEco*.

35 La méthode suivie est conforme à celle suivie pour la réalisation de la banque "*MboI*" décrite ci-dessus. De plus fortes quantités d'ADN méningococciques ont cependant été utilisées pour la première itération

5 d'hybridation soustractive afin de compenser le plus grand nombre de fragments de faibles poids moléculaires produits par *Tsp509I*. Pour la première itération, 400 ng de fragments d'ADN Nm et, dans la seconde, 25 ng de fragments Nm sont soumis à hybridation soustractive avec 40 µg d'ADN Ng cisaillé de manière aléatoire.

10 Pour la réalisation de cette banque "Tsp509I", à titre de contrôle, une troisième itération d'hybridation soustractive est réalisée en utilisant 40 µg d'ADN Ng cisaillé et 0,2 ng de fragments Nm résultant d'une digestion par *Tsp509I* et d'une re-ligature aux adaptateurs NEco des fragments obtenus à l'issue de la seconde itération.

15 Spécificité - Comme décrit pour la banque précédente, le produit issu de la deuxième itération de la méthode CDA est marqué et utilisé comme sonde pour un panel de souches de *Neisseria*.

20 La figure 1A illustre l'hybridation Southern blot des produits de la seconde itération de la méthode CDA avec l'ADN digéré par *ClaI* de : Nm en piste a, de Ng MS11 en piste b, de Nm 8013 en piste c, de Ng 403 en piste d, de Nm 1121 en piste e, de Ng 6934 en piste f, de Nm 1912 en piste g, de Ng WI (souche DGI) en piste h, de Nm 7972 en piste i, de Nl 8064 en piste j, de Nc 32165 en piste 25 k, de Nm 8216 en piste l.

Contrairement à la forte réactivité observée avec toutes les souches Nm, une faible, ou aucune réactivité, est observée avec les souches Ng, Nl et Nc.

30 La spécificité de la banque a été étudiée plus avant en faisant réagir des transferts sur membrane (Southern blots) des produits issus de chacune des trois itérations de la méthode CDA avec des sondes correspondant à *pilC1* et *ppk*. Ces deux gènes sont communs à Nm et Ng.

35 La figure 1B représente un gel d'agarose après électrophorèse des chromosomes de Nm Z2491 et Ng Ms11,

digérés avec *Tsp509* et des produits issus de chacune des itérations de la méthode CDA.

En piste a, a été déposé 1 µg du chromosome de Nm, en piste b 1 µg de celui de Ng, en piste c 0,15 µg des 5 produits issus de la première itération CDA, en piste d 0,1 µg de ceux de la seconde itération, en piste e 0,05 µg de la troisième itération, MW représentant les marqueurs de taille moléculaire.

Les figures 1C et 1D représentent des gels réalisés 10 comme décrits en figure 1B après transfert sur membrane (Southern blots) et hybridation avec *pilC1* (figure 1C) et *ppk* (figure 1D).

A l'issue de la seconde intégration de la méthode CDA, les séquences correspondant aux gènes *pilC1* et *ppk* 15 sont complètement exclues de la banque.

Exhaustivité - L'exhaustivité de la banque a été examinée en faisant réagir les produits issus de l'hybridation soustractive avec des sondes correspondant à trois gènes Nm-spécifiques (*frp*, *rotamase* et *opc*).

20 Ces sondes Nm-spécifiques réagissent avec les produits d'amplification issus de la première et de la seconde intégration d'hybridation soustractive.

Les figures 1E, 1F et 1G représentent des gels réalisés comme décrits en figure 1B après transfert sur 25 membrane (Southern blots) et hybridation avec *frpA* (figure 1E), *rotamase* (figure 1F) et *opc* (figure 1G).

Une troisième intégration d'hybridation soustractive conduit cependant à la perte de séquences Nm-spécifiques car les fragments réagissant avec les gènes *rotamase* et 30 *opc* sont absents de cette troisième intégration.

En considérant l'ensemble de ces données, il résulte que les produits issus de la seconde intégration de la méthode CDA sont Nm-spécifiques et constituent également une banque exhaustive des séquences Nm-spécifiques.

Les produits issus de cette deuxième itération sont clonés au niveau du site EcoRI du plasmide pBluescript.

La banque produite par *Tsp509I* est plus exhaustive que la banque produite par *MboI*, comme les 5 considérations théoriques basées sur la production enzymatique de plus petits fragments de restriction le supposaient.

Selon cet aspect, il faut aussi noter que la banque *Tsp509I* est moins redondante que la banque *MboI* c'est-à-dire qu'elle comprend moins de duplication de clones. 10 86% des clones de la banque *Tsp509I* correspondent à des séquences distinctes alors que seulement 43% des clones correspondent à des séquences distinctes dans la banque *MboI* (données non présentées).

15 La banque produite par *Tsp509I* constitue donc une source de clones Nm-spécifiques.

Exemple 2 : Analyse des clones des banques soustractives

20 Clonage et séquençage des ADN Nm-spécifiques

Les ADN des banques soustractives sont clonés au niveau du site *BamHI* (banque *MboI*) ou *EcoRI* (banque *Tsp509I*) du plasmide pBluescript, puis transformés dans DH5 α de *E. coli*. Les inserts sont amplifiés par PCR réalisée sur les colonies transformées en utilisant les 25 amorces M13-50 et M13-40, cette dernière amorce étant biotinylée à son extrémité 5'.

Le séquençage a été réalisé sur chaque produit PCR après séparation des brins biotinylés et non-biotinylés 30 en utilisant le système Dynabeads M-280 à streptavidine (Dynal, Oslo). Les séquences sont ciblées selon leurs homologies avec des séquences précédemment publiées en utilisant les programmes informatiques Blastn et Blastx (NCBI, USA et Fasta).

Les produits PCR issus des colonies de bactéries transformées, après utilisation des amorces M13-40 et M13-50 comme décrit ci-dessus, ont été marqués par incorporation avec amorçage aléatoire de α - ^{32}P -dCTP et ont été utilisés comme sonde pour les transferts sur membrane de l'ADN chromosomique digéré par *Cla*I des souches Nm Z2491 et Ng MS11, comme décrit ci-dessus afin de vérifier leur spécificité.

10 **Cartographie des clones sur le chromosome de la souche Nm Z2491.**

On rapporte les résultats des études effectuées avec 17 clones de la banque "MboI" (désignés par la lettre B) et 16 clones de la banque "Tsp5091" (désignés par la 15 lettre E), chacun de ces clones présentant une séquence unique et sans contrepartie chez Ng.

20 Les positions des séquences d'ADN correspondant aux produits Nm-spécifiques clonés ont été déterminées par rapport à la carte publiée du chromosome de Nm Z2491 (Dempsey et al. 1995, *J. Bacteriol.* 177, 6390-6400) et à l'aide de transferts sur membranes (Southern blots) de gels d'agarose ayant été soumis à électrophorèse à champ pulsé (PFGE).

25 Les clones Nm-spécifiques sont utilisés comme sondes pour une hybridation sur membranes (Southern blots) de l'ADN de Nm Z2491 digéré avec des enzymes à rares sites de coupure, à savoir *Pac*I, *Pme*I, *Sgf*I, *Bgl*III, *Spe*I *Nhe*I que *Sgf*I.

30 Les gels (20 x 20 cm) étaient des gels à 1% d'agarose dans un tampon TBE 0,5X et ont été soumis à électrophorèse à 6 V/cm pendant 36 heures selon des périodes de pulsation variant de manière linéaire entre 5 et 35 secondes.

35 Les hybridations sur membrane (Southern blots) ont été réalisés comme décrit précédemment.

Les résultats obtenus sont rapportés sur la figure 2 : la réactivité a été localisée par comparaison avec les positions des fragments de taille correspondante sur la carte publiée. Les positions de l'ensemble des 5 marqueurs génétiques cartographiés par Dempsey et al (précédemment cité) sont visualisées à l'aide de points sur la carte linéaire chromosomique. Les gènes Nm-spécifiques précédemment divulgués sont marqués d'un astérisque. Les deux loci appelés "frp" correspondent aux 10 gènes *frpA* et *frpC*. Les locis "pilC" correspondent aux gènes *pilC1* et *pilC2* qui sont des paires de gènes homologues et qui ne sont pas distingués sur la carte. La précision des positions des clones Nm-spécifiques de 15 l'invention dépend des chevauchements des fragments de restriction réactifs. En moyenne, la position est de +/- 20 kb.

Cette cartographie révèle une distribution non aléatoire des séquences Nm-spécifiques. La majorité des 20 séquences Nm-spécifiques appartiennent à trois groupes distincts. Un de ces groupes (région 1) correspond à la position de gènes relatifs à la capsule précédemment décrits.

On distingue :

- E109, E138, B230 et B323 comme étant la région 1,
- 25 - B322, B220, B108, B132, B233, B328, E139, E145 et B101 comme étant la région 2, et
- B306, E114, E115, E124, E146, E120, E107, E137 et E142 comme étant la région 3.

30 63% des séquences identifiées comme spécifiques des méningocoques sont localisées à l'intérieur de ces trois régions distinctes.

Ce regroupement contraste avec la distribution de gènes Nm-spécifiques précédemment divulgués (*frpA*, *frpC*, *porA*, *opc* et la région relative à la capsule).

Cet art antérieur suggérait en effet que les gènes Nm-spécifiques étaient à l'exception des gènes fonctionnellement relatifs à la capsule, dispersés le long du chromosome.

5 La cartographie des séquences Nm-spécifiques sur le chromosome conduit à un résultat inattendu en regard de l'art antérieur.

10 La majorité des différences génétiques entre les souches meningocciale et gonococciale testées sont regroupées en trois régions distinctes.

La région 1 regroupe des gènes relatifs à la capsule des meningococci.

15 La fonction des gènes des autres régions n'est pas connue mais des homologies avec des séquences publiées (tableau 1) suggèrent des similarités entre certains gènes de la région 3 et les protéines transposases et de régulation de bactériophages. Aucun virus meningococcal n'a été caractérisé et il est tentant d'imaginer que ces séquences soient d'origine phagique. De manière 20 intéressante, le génome de *H. influenzae* contient également une séquence homologue à celle de la protéine de régulation *Ner* du phage *Mu* mais on ne sait pas s'il s'agit d'un gène fonctionnel.

25 Le clone B208 présente une forte homologie (48% d'identité, 91% d'homologie pour 33 acides aminés) avec un clone des régions conservées (domaine III) dans la classe des protéines qui se lient aux sidérophores ferriques TonB-dépendants.

30 La proximité de ce clone avec les gènes Nm-spécifiques *porA* et les gènes régulés par le fer *frp*, et en particulier la possibilité qu'il s'agisse d'une protéine récepteur Nm-spécifique exposée sur la membrane externe font de lui un bon candidat pour de plus amples recherches.

Le clone B339 correspond à la séquence d'insertion Nm-spécifique IS1106.

La faible homologie entre le clone B134 et la séquence d'insertion d'Aeromonas, ainsi que la présence en copies multiples du clone B134 parmi des souches variées de Nm, suggèrent qu'il pourrait représenter un nouveau type de séquence d'insertion Nm-spécifique.

La possibilité pour que les régions contenant les clones Nm-spécifiques puissent correspondre à des îlots de pathogénicité comme précédemment décrit pour *E. coli* et *Y. pestis* est d'un intérêt particulier.

Les clones isolés dans cette invention vont permettre de mieux comprendre la pertinence des régions Nm-spécifiques en permettant le clonage et le séquençage de fragments chromosomiques plus grands et directement par leur utilisation pour des mutations de loci.

Enfin, la détection des gènes meningococcus-spécifiques, éventuellement impliqués dans la pathogénicité de l'organisme, permet de cibler des antigènes appropriés utilisables dans un vaccin anti-meningococcique.

L'efficacité et la rapidité de la méthode selon l'invention permettent son utilisation dans un grand nombre de situations pour lesquelles les différences génétiques responsables d'un phénotype particulier à un de 2 pathogènes proches sont recherchées.

Etude de la réactivité des clones des régions 1, 2 et 3 vis-à-vis d'un panel de souches de *Neisseria*

Les produits PCR correspondant aux inserts de chacun des clones ont été rassemblés et utilisés comme sondes d'hybridation sur membranes (Southern blots) pour un panel de souches de Nm, de Ng, de Nl et de Nc.

Les régions 1 et 2 produisent un nombre limité de bandes pour chacun des méningocoques. Cela suggère que

ces régions sont à la fois Nm-spécifiques et communes à tous les méningocoques.

La figure 3 illustre la réactivité des clones des régions 1, 2 et 3 envers un panel de souches 5 neisseriales. Les clones des régions 1 (figure 3A), 2 (figure 3B) et 3 (figure 3C) pris ensemble ont été utilisés comme sondes envers un panel de meningococci, gonococci et envers une souche de N1 et de Nc.

Les pistes sont les suivantes : ADN de Nm Z2491 en 10 piste a, de Ng MS11 en piste b, de Nm 8013, en piste c, de Ng 403 en piste d, de Nm 1121 en piste e, de Ng 6934 en piste f, de Nm 1912 en piste g, de Ng WI (souche DGI) en piste h, de Nm 7972 en piste i, de N1 8064 en piste j, de Nc 32165 en piste k, de Nm 8216 en piste l.

15 De manière remarquable, la région 3 ne présente de réactivité qu'avec les meningococci de sérogroupe A. Cette région 3 est donc spécifique d'un sous-groupe de Nm.

Une comparaison avec des séquences connues dans les 20 banques de données a été réalisée afin d'évaluer les possibles fonctions des régions clonées.

Le tableau 1 qui suit donne les positions des clones spécifiques sur la carte chromosomique et les homologies avec des séquences connues.

TABLEAU 1 - Position des clones spécifiques sur la carte chromosomique et homologies avec des séquences communes

Nom du Clon*	Taille de l'insert	Pac	Fragments réactifs				Position sur Z249]	Homologies des séquences protéiques
			Pme	Bgl	Spc	Nhe		
B305	259	18-20	15-17	22-23	18	11-13	2	
B333	235	15-17	22-23	18	11-13	2	λ736	
E109 [†]	211	6-7	11-15	10	11-13	2	<i>mafA</i> <i>craA</i>	protéine LipB <i>N. meningitidis</i> (3 x 10 ⁻²⁴)
E138 [†]	315	1	6-7	11-15	10	11-13	2	<i>mafA</i> <i>craA</i>
B230 [†]	356	1-3	6-7	1	10	11-13	2	protéine LipB <i>N. meningitidis</i> (4 x 10 ⁻²⁵)
B323 [†]	363	1	6-7	1	10	11-13	2	protéine KpsC <i>E. coli</i> (3 x 10 ⁻²⁵)
B322 [‡]	210		2	16-18	6	1	5	protéine CtrB <i>N. meningitidis</i> (2 x 10 ⁻²⁴)
B220 [‡]	341		2	16-18	6	≥18	5	<i>HyB S. marcescens</i> (4 x 10 ⁻²⁵)
B108 [‡]	275		2	19-21	6	≥18	5	<i>pil()λ740</i>
B132 [‡]	411	2	2	19-21	6	≥18	5	<i>pil()λ740</i>
B233 [‡]	164	1-3	2	19-21	6	≥18	5	<i>pil()λ740</i>
B328 [‡]	256	1-3	2	22-23	6	≥18	5	<i>pil()λ740</i>
E139 [‡]	324	2	2	19-21	6	≥18	5	<i>pil()λ740</i>
E145 [‡]	343	2	2	19-21	6	≥18	5	<i>pil()λ740</i>
B101 [‡]	254	≥20	2	19-21	6	≥18	5	<i>pil()λ740</i>
E103q	334		2	11-15	3-5	10	3	
B326 [‡]	314		2	11-15	3-4	10	3	
B326 (faible réactivité)			5	6	16	2	1	<i>argF</i>
B342	167		2	19	3-4	6-7	3	<i>lga</i>
E136	249		2	7	1	3	3	<i>kpa</i>

	B208	177	1	2	3-4	2	1	<i>parA</i>	Récepteur de la pirocheline FcIII <i>P. aeruginosa</i> (5×10^{-4})
= B306 ^a	219	11	5	11-12	5	2	4	parC	
E114 ^a	227	11	5	11-12	5	2	4	parC	
E115 ^a	251		5	11-15	5	2	4	parC	
E124 ^a	208		5	11-12	5	2	4	parC	
E146 ^a	146		5	11-15	5		4	parC	
E120 ^a	263		5	3-4	5	16	4	opαB	
E107 ^a	248	11	14-17	3-4	5	16	4	opαB	
E137 ^a	274		14-17	3-4	5	16	4	opαB	Transposase Bacteriophage D3112 (6×10^{-4})
E142 ^a	230		14-17	3-4	5	16	4	opαB	Protéine Ner-Like <i>H. influenzae</i> (6×10^{-3}) Protéine se liant à l'ADN Ner. Phage mu (3×10^{-18})
E116	379	5-7	11-13	3-4	2	6-7	8	λ375	
B313	436	9	9	3-4	13-14	5	2	λ611	
B341	201	8-10	9	3-4	13-14	5	2	λ611	
E102	238		11-13	3-4	19	5	2	λ601	Protéine hypothétique III1730 <i>H. influenzae</i> (7×10^{-24}) transposase ISAS2.
B134	428				multiple				<i>Aeromonas</i>
B339	259				multiple				<i>salmonicida</i> (5×10^{-3}) transposase IS 1106 <i>N. meningitidis</i> (6×10^{-3})

FEUILLE DE REMPLACEMENT (REGLE 26)

Entre Parenthèses figure la signification des homonymies françaises, celle que donne par le procédé d'analyse

Les clones marqués de l'exposant 1-2 ou 3 appartiennent aux régions 1-2 ou 3, respectivement du chromosome de *X. meningitis* 22491.

de chevauchement #1 B216 présente également une faible réactivité dans la région de *new F*.

(q) Le clone E103 contient un site *Typ509* qui peut donc contenir deux inserts, cependant, comme il ne réagit qu'avec un seul fragment (*Gal* (Oks) du chromosome de *N. meningitidis* 22:49) et l'occupe qu'une position sur la carte, ce clone est inclus ici.

On peut voir, tout d'abord, que les clones de la région 1 correspondent tous aux gènes impliqués dans la biosynthèse de la capsule. Ces gènes ont été précédemment étudiés parmi les Nm de sérogroupe B (Frosch et al. 1989, 5 Proc. Natl. Acad. Sci. USA 86, 1669-1673 et Frosch et Muller 1993, Mol. Microbiol. 8 483-493).

A l'exception d'une faible homologie avec l'activateur de hémolysine de *Serratia marcescens*, les clones de la région 2 ne présentent aucune homologie 10 significative avec les séquences publiées, que ce soit au niveau de l'ADN ou des protéines.

Deux des clones de la région 3 présentent d'intéressantes homologies avec des protéines qui se lient à l'ADN, en particulier les protéines de régulation 15 et les protéines transposases de bacteriophages.

Le clone B208 présente une forte homologie avec une des régions conservées dans une classe de récepteurs (sidérophore ferrique TonB-dependant).

Les clones B134 et B339 s'hybrident avec de 20 nombreuses régions du chromosome (au moins 5 et au moins 8, respectivement).

Les données concernant les séquences montrent que le clone B339 correspond à la séquence d'insertion Nm-spécifique S1106.

25 La traduction du clone B143 présente une homologie limitée avec la transposase d'une séquence d'insertion *Aeromonas* (SAS2)(Gustafson et al. 1994, J. Mol. Biol. 237, 452-463). Nous avons pu démontrer par transfert sur membrane (Southern blots) que ce clone est une entité Nm-spécifique présente en multiples copies dans les 30 chromosomes de chaque meningocoque du panel testé.

Les autres clones ne présentent pas d'homologie significative avec les séquences neisseriales publiées ni d'ailleurs avec aucune séquence publiée. Ces clones 35 constituent donc, avec la majorité des autres clones

isolés, une banque de loci Nm-spécifiques totalement nouveaux.

Exemple 3: Etude de la région 2 du chromosome de Nm

5

. Détermination et caractérisation de la séquence de la région 2

On procède à une amplification par PCR avec de l'ADN chromosomique de la souche Z2491 de sérogroupe A, sous-groupe IV-1, en utilisant des amorces d'oligonucléotides élaborées à partir de chacune des séquences de clones de la région 2, selon de nombreuses combinaisons différentes. On séquence les produits de la PCR qui se chevauchent à partir des 2 brins en utilisant la technique de terminaison de chaîne et le séquençage automatisé (ABI 373 ou 377).

Pour prolonger la séquence au-delà des limites des clones disponibles, on clone des fragments partiels SauIIIA de 15 kb, de la souche Z2491, dans Lambda DASH-II (Stratagène).

On identifie les phages contenant les inserts chevauchant la région 2 par hybridation avec comme sondes des clones de cette région. L'ADN inséré est séquencé à partir des extrémités des inserts et ces séquences sont utilisées pour élaborer de nouvelles amorces qui serviront à amplifier directement l'ADN chromosomique et non l'ADN phagique.

On obtient une amplification de l'ADN chromosomique en utilisant ces nouvelles amorces et celles de la séquence précédemment disponible.

Ces produits PCR sont également séquencés à partir des 2 brins, ce qui conduit à une séquence complète de 15620 pb (SEQ ID N°36). On analyse les cadres de lecture de cette séquence qui commencent par ATG ou GTG et qui sont caractérisés par un indice d'usage de codons élevés.

Cette analyse révèle 7 COLs de ce type qui remplissent la plus grande partie de la séquence de 15620pb. Les positions de ces COLs sont les suivantes:

COL-1: 1330 à 2970 (SEQ ID N°37); COL-2: 3083 à 9025 (SEQ ID N°38); COL-3: 9044 à 9472 (SEQ ID N°39); COL-4: 10127 à 12118 (SEQ ID N°40); COL-5: 12118 à 5 12603 (SEQ ID N°41); COL-6: 12794 à 13063 (SEQ ID N°43); COL-7: 13297 à 14235 (SEQ ID N°44); et COL-8: 14241 à 15173 (SEQ ID N°45).

Le COL-4 commence avec le codon GTG et chevauche un COL légèrement plus petit (SEQ ID N°41) dans le même cadre de lecture (9620-12118, cadre 2) et qui commence par le codon ATG.

10 COL-4 code pour une protéine qui présente des homologies structurelles avec une famille de polypeptides comprenant les pyocines (*Pseudomonas aeruginosa*), collcines et intimines (*Escherichia coli*) qui sont des toxines bactéricides (pyocines, collcines) ou des protéines de surfaces impliquées dans l'adhésion des bactéries aux protéines eucaryotes. Le COL-7 encode une protéine dont la séquence contient une 15 région potentiellement transmembranaire, et qui présente des homologies structurelles avec des protéines périplasmiques ou insérées dans la membrane externe des bactéries. Les homologies structurelles de COL-4 et COL-7 ont été identifiées à l'aide du programme PropSearch.

20 La recherche de séquences homologues aux autres COL dans GenBank à l'aide du programme BLAST a révélé une homologie entre les régions N-terminales de COL-2 et l'hémagglutinine filamentuse B de *Bordetella pertussis* (43% de similarité, 36% d'identité sur 352 acides aminés) et entre COL-1 et la protéine fhaCde *Bordetella pertussis* (35% de similarité, 27% d'identité sur 401 acides aminés). COL-1 et COL-2 sont des gènes voisins dans la souche Z2491 et l'hémagglutinine filamentuse B de 25 *Bordetella pertussis* et fhaC sont des gènes voisins dans *Bordetella pertussis*, ce qui renforce la probabilité que ces homologies reflètent des homologies fonctionnelles.

Confirmation de la spécificité de la région 2 vis-à-vis de Nm

On effectue des Southern blots en utilisant des sondes d'ADN obtenues par amplification par PCR de différentes parties de la région 2 en utilisant des amorces 30 oligonucléotidiques élaborées à partir de séquences de clones de la région 2.

On a représenté sur la figure 4 la position approximative de ces oligonucléotides.

Il s'agit, dans une moitié de COL-1, des oligonucléotides appelés R2001 (SEQ ID N°46) et R2002 (SEQ ID N°47), dans une moitié de COL-1+la majeure partie de COL-2, des oligonucléotides b332a (SEQ ID N°48), e139a (SEQ ID N°49), b132a (SEQ ID N°50) et b233b (SEQ ID N°51), et dans 1/3 de COL-4+ COL-5 à 7, des oligonucléotides e145a (SEQ ID N°52) et b101a (SEQ ID N°53).

Les trois Southern sont réalisés dans les 10 conditions d'hybridation suivantes:

16 h à 65°C,
NaPO₄ 0,5M, pH 7,2
EDTA-Na 0,001M
1% de dodécylsulfate de sodium.

15

Pour le lavage, on chauffe 10 min à 65°C et on utilise NaPO₄ 0,5M, pH 7,2; EDTA-Na 0,001M, 1% de dodécylsulfate de sodium.

Les figures 5, 6 et 7 représentent respectivement 20 les Southern blots obtenus avec chacune des parties de COL mentionnées plus haut.

Les 14 pistes correspondent respectivement, dans chacun des Southern, à

- 1: MS11 (Ng)
- 25 2: 403 (Ng)
- 3: FA1090 (Ng)
- 4: W1 (Ng)
- 5: 6493 (Ng)
- 6: marqueur (lambda hindIII)
- 30 7: Z2491 (Nm, gpA)
- 8: 7972 (Nm gpA)
- 9: 8013 (Nm, gpC)
- 10: 1121 (Nm non groupable)
- 11: 1912 (Nm, gpB)
- 35 13: 32165 (Nc)

14: 8064 (N1).

Etant donné qu' un panel de souches de *Neisseria* est utilisé dans ces expériences et que chaque puits est chargé avec une quantité similaire d'ADN digéré, ces 3 Southern blots montrent clairement que les séquences correspondant à la région 2 sont trouvées dans tous les méningoccoques testés et qu'il n'existe pas dans le génome de Ng des souches testées de séquences homologues significatives.

Exemple 4: Identification de régions du génome de Nm absentes de N1 et communes avec Ng

On opère selon la technique décrite dans l'exemple 1, mais on utilise l'ADN chromosomique d'une souche de Nm (Z2491) et de 2 souches de N1 (collection XN) dont on mélange les ADN à parts égales.

On effectue 2 soustractions en utilisant les séries d'amorces R et J. Trois banques différentes sont ainsi réalisées.

Deux banques, appelées Bam et Eco, sont respectivement obtenues par digestion de l'ADN chromosomique de Nm Z2491 par *MboI* et *Tsp509I*; une troisième banque, appelée Cla, qui résulte de la digestion de l'ADN chromosomique de Nm par *MspI*, est obtenue en utilisant le jeu d'amorces RMsp10, RMsp24, JMsp10 et JMsp24. L'ensemble des amorces utilisées est donné dans le tableau 2 suivant.

Tableau 2

5 Adaptateurs pour banques différentielles

10 ADN chromosomique digéré par Clonage dans
pBluescript par

<i>MboI</i>	→	<i>BamHI</i>
<i>Tsp509I</i>	→	<i>EcoRI</i>
<i>MspI</i>	→	<i>ClaI</i>

15

Premier tour de soustraction

20 RBam12 : 3' AGTGGCTCCTAG 5' (SEQ ID N°54)
 RBam24 : 5' AGCACTCTCCAGCCTCTCACCGAG 3' (SEQ ID N°55)

25 REco12 : AGTGGCTCTTAA (SEQ ID N°56)
 RBam24 : 5' AGCACTCTCCAGCCTCTCACCGAG 3' (SEQ ID N°55)
 (REco 24 = RBam 24)

30 RMsp10 : AGTGGCTGGC (SEQ ID N°57)
 RMsp24 : 5' AGCACTCTCCAGCCTCTCACCGAC 3' (SEQ ID N°58)

Deuxième tour de scission

35 Jbam12 : 3' GTACTTGCCTAG 5' (SEQ ID N°59)
 JBam24 : 5' ACCGACGTCGACTATCCATGAACG 3' (SEQ ID N°60)
 JEcol2 : GTACTTGCCTAA (SEQ ID N°61)
 JBam24 : 5' ACCGACGTCGACTATCCATGAACG 3' (SEQ ID N°60)
 40 (JEco 24 = JBam 24)
 JMsp10 : GTACTTGGGC (SEQ ID N°62)
 JMsp24 : 5' ACCGACGTCGACTATCCATGAACC 3' (SEQ ID N°63)

Après 2 soustractions, on marque la totalité du produit de chaque amplification et on l'utilise comme sonde.

5 On effectue un contrôle des banques soustractive par Southern blot sur un panel de 12 souches de *Neisseria* (ADN chromosomique coupé par *ClaI*). Les conditions d'hybridation sont identiques à celles données dans l'exemple 1.

10 Ces Southern blots sont donnés sur les figures 8A à 8C, qui sont respectivement relatives à la banque *MboI/BamHI*, à la banque *MspI/ClaI* et à la banque *Tsp509I/EcoRI*.

Les 12 pistes correspondent respectivement à :

- 15 1: Nm Z2491 (groupe A)
- 2: N1 8064
- 3: Nm 8216 (groupe B)
- 4: N1 9764
- 5: Nm 8013 (groupe C)
- 20 6: Ng MS11
- 7: Nm 1912 (groupe A)
- 8: Ng 4C1
- 9: Nm 1121 (non groupable)
- 10: Ng FA1090
- 25 11: Nc 32165
- 12: Nm 7972 (groupe A).

30 L'examen des Southern blots montre que les séquences contenues dans chaque banque sont spécifiques de Nm et ne sont pas trouvées chez N1. De plus, la réactivité observée avec les souches de Ng suggère que certaines de ces séquences sont présentes chez Ng.

35 Chacune de ces banques a ensuite été clonée dans pBluescript au site *BamHI* pour Bam, ou *EcoRI* pour Eco, ou *ClaI* pour Cla. Afin de confirmer la spécificité des

clones vis-à-vis du génome de Nm, on a procédé à une restriction des clones individuels et à leur radiomarquage. Les clones montrant à la fois une réactivité pour Nm et Ng ont été conservés pour des études ultérieures. Ces clones sont représentés sur les figures 9, 10 et 11, qui donnent les profils, vis-à-vis de Nm, Nl et Ng, de 5 clones de la banque Bam (figure 9), de 16 clones de la banque Eco (figure 10), et de 13 clones de la banque Cla (figure 11).

Ces clones ont été séquencés en utilisant des amorces universelles et inverses. Il s'agit

- des clones Bam

B11 partiel de 140 pb (SEQ ID N° 66), B13 partiel estimé à 425 pb (SEQ ID N° 67), B26 de 181 pb (SEQ ID N° 68), B33 de 307 pb (SEQ ID N° 69), B40 de 243 pb (SEQ ID N° 70),

- des clones Cla

C16 de 280 pb (SEQ ID N° 72), C20 partiel estimé à 365 pb (SEQ ID N° 73), C24 partiel estimé à 645 pb (SEQ ID N° 74), C29 partiel estimé à 245 pb (SEQ ID N° 75), C34 de 381 pb (SEQ ID N° 76), C40 de 269 pb (SEQ ID N° 77), C42 de 203 pb (SEQ ID N° 78), C43 de 229 pb (SEQ ID N° 79), C45 de 206 pb (SEQ ID N° 80), C47 de 224 pb (SEQ ID N° 81), C62 de 212 pb (SEQ ID N° 82), et C130 (5'...) estimé à 900 pb (SEQ ID N° 83), et

- des clones Eco

E2 de 308 pb (SEQ ID N° 84), E5 partiel, estimé à 170 pb (SEQ ID N° 85), E22 partiel estimé à 300 pb (SEQ ID N° 86), E23 de 273 pb (SEQ ID N° 87), E24 de 271 pb (SEQ ID N° 88), E29 de 268 pb (SEQ ID N° 89), E33 partiel, estimé à 275 pb (SEQ ID N° 90), E34 partiel, estimé à 365 pb (SEQ ID N° 91), E45 de 260 pb (SEQ ID N° 92), E59 estimation supérieure à 380 pb (SEQ ID N° 93), E78 de 308 pb (SEQ ID N° 94), E85 de 286 pb (SEQ ID N° 95), E87 de 238 pb (SEQ ID N° 96), E94 partiel, supérieur à 320 pb

(SEQ ID N° 97), E103 partiel, supérieur à 320 pb (SEQ ID N° 98) et E110 de 217 pb (SEQ ID N° 99).

La cartographie de chaque clone a été effectuée sur le chromosome de Nm 22491 en opérant comme décrit dans 5 l'exemple 2. Les résultats obtenus sont donnés sur la partie droite de la figure 2. On constate que ces clones correspondent aux régions appelées 4 et 5. Ces régions sont donc constituées de séquences présentes à la fois chez Nm et chez Ng, mais non trouvées chez N1. Il est 10 donc considéré qu'il s'agit de séquences codant pour des facteurs de virulence responsables de la colonisation initiale et de la pénétration de la muqueuse. La région 4 est localisée entre argF et regF sur le chromosome de Nm 2491 et la région 5 entre le marqueur lambda 375 et penA. 15 Cette région contient vraisemblablement des séquences codant pour un variant Opa et une protéine liant la transferrine.

Une comparaison avec les séquences connues dans les 20 banques de données a montré que dans la région 4 seul le clone C130 présente une homologie, à savoir avec *MspI* méthylase. Dans la région 5, aucune homologie avec des séquences connues n'a été trouvée avec les clones C8, E2, B40, C45, E23 et E103. Pour les autres clones, les 25 homologies sont les suivantes :

B11 arginine décarboxylase SpeA; C29 arginine décarboxylase SpeA; C62 oxoglutarate/malate transporteur; repetitive DNA element; E34 élément répétitif d'ADN ; E94 endopeptidase MepA murine ; C47 citrate synthase PrpC; E78 citrate synthase PrpC 30

Exemple 5 : Mise en évidence de la présence d'une ou plusieurs souches de *Neisseria meningitidis* dans un échantillon biologique.

Un échantillon biologique de type liquide céphalo- 35 rachidien, urine, sang, salive est prélevé.

Après filtration et extraction, les ADN présents dans cet échantillon sont soumis à électrophorèse sur gel et transférés sur membrane par Southern blotting.

Une sonde nucléotidique constituée par le marquage 5 au ³²P de la SEQ ID n°5 est incubée avec cette membrane de transfert.

Après antoradiographie, la présence de bande(s) réactive(s) permet de diagnostiquer la présence de *Neisseria meningitidis* dans l'échantillon.

10

Exemple 6 : Composition vaccinale incluant dans son spectre une prophylaxie à visée anti-méningococcique et destinée à prévenir toute forme d'infection par *Neisseria meningitidis*.

15

Le peptide codé par une séquence incluant la SEQ ID n°10 est conjugué à une toxine.

Ce peptide conjugué est alors ajouté à une composition comportant le vaccin anti-*Haemophilus* et anti-pneumocoque, ou tout autre vaccin de l'enfance.

20

La composition résultante peut, après avoir été rendue stérile, être injectée par voie parentérale, sous-cutanée ou intramusculaire.

Cette même composition peut également être pulvérisée au niveau des muqueuses à l'aide d'un spray.

25

46
LISTE DE SEQUENCES

(1) INFORMATIONS GENERALES:

(i) DEPOSANT:

- (A) NOM: I.N.S.E.R.M
- (B) RUE: 101, rue de Tolbiac
- (C) VILLE: PARIS CEDEX 13
- (E) PAYS: FRANCE
- (F) CODE POSTAL: 75654

(ii) TITRE DE L' INVENTION: ADN, protéines et peptides spécifiques des bactéries de l'espèce *Neisseria meningitidis*, leurs procédés d'obtention et leurs applications biologiques.

(iii) NOMBRE DE SEQUENCES: 99

(iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk
- (B) ORDINATEUR: IBM PC compatible
- (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
- (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (CEB)

(2) INFORMATIONS POUR LA SEQ ID NO: 1:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 257 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

GATCCGCTGC CGGCAGACGA ATATCAAGAC ATCTTCGATT TTATGAAACA GTATGACTTG

60

TCTTACCCGT ATGAATATCT GCAGGATTGG ATAGATTACT ATACGTTCAA AACCGATAAG	120
CTGGTATTTG GTAACCGCAA GCGAGAGTGA GCGGTAAAAC TCTGAGCTCC TGTGTTATAG	180
ATTACAACCT TAGGCGCTT TAAAGCTGAA AGATTTCTGA AAGCTATAAA TTGAAGCCCT	240
TCCACAGTAC ATAGATC	257

(2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 276 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

GATCATGTTTC AAATAGATAG GCATGGGAAG CTGCAGCTCT AACGTCCATG AAAATATGTT	60
GCATAGCTGC AAGCGGAACG CCTTTTCTTT CATCTACATA ATCTATAGAG TCAAGGCAAC	120
CGCTATTGAA ATTAGCAGTA TTGCCTATGA TTACATTAGT AATATGCTCA TACCATTTTT	180
GGGTGGTCAT CATAATTGTGC CCCATTGTGA TCTCCTTATA TTGGTTTAG AAGGAACCTT	240
GACAGGAAGA ATAACGGCCT TACCTGTTTG ACGATC	276

(2) INFORMATIONS POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 428 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*
 (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

GATCTGGTGG TGTGTCACA GGTAGGCGCA TACTTGTTCG GGACTGAGTT TCGGGCGGAT 60
 AAGGGTGTGAA ATGTGCTGAA TCAGCTGCGA ATCGAGCTTA TAGGGTTGTC GCTTACGCTG 120
 TTTGATAGTC CGGCTTTGCC GCTGGGCTTT TTCGGCGCTG TATTGCTGCC CTTGGGTGCC 180
 GTGCCGCTG ATTTCGCGGC TGATGGTGCT TTTGTGGCGG TTAAGCTGTT TGGCGATTTG 240
 GGTGACGGTG CAGTGGCGGG ACAGGTATTG GATGTGGTAT CGTTCGCCTT GGGTCAGTTG 300
 CGTGTAGCTC ATGGCAAATCT TTCTTGCAAGG AAAGGCCGTA TGCTACCGCA TACTGGCCTT 360
 TTTCTGTTAG GGAAAGTTGC ACTTCAAATG CGAATCCGCC GACCTCTTTC AGTTACAGCA 420
 GCTTGATC 428

(2) INFORMATIONS POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 390 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*
 (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

GATCCTGCAT TGACATCGGC CTTGGCTGTC AGGGTATTGT GACCGGTAAA GTGGCATTA 60
 CCGTTGGCCA ATAAGGATAC ATGACCGTCT GCAGAAACAG CATGAAGGCC GTCTGAAACG 120
 ATATTGCCCT GCAATGCGGT GGTTTCGAGA GCCTTGGCTG CGTTCAAGCTT GGTATTGCGA 180
 AGCTGAATAT TGCCTTGGC TGCCTGAATG TGCAGATTAC CCGAGTTGGT ACGCAGATTG 240

GTATTGGTAA CATTCAAGCAA GCCTGCCTCC ACACCCATGT CTTTTGAGGC AGTGAGGGTT	300
TTACTGGTGC CGGTAATATG GGCAAGCTTA TCCGATTTCA AATGGATGCT GGCCGGCAGA	360
CAAATCTTTA TCAACATTCA AATTCAAGATC	390

(2) INFORMATIONS POUR LA SEQ ID NO: 5:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 177 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

GATCAGATTG GTGAAGACGG TATTACCGTC AATGTTGCAG GCCGTTCCGG ATATACGGCG	60
AAAATCGACG TGTCTCCGAG TACCGATTG GCGGTTTATG GCCATATTGA AGTTGTACGG	120
GGTGCAACGG GGTTGACCCA ATCCAATTCA GAGCCGGGTG GAACCGTCAA TTTGATC	177

(2) INFORMATIONS POUR LA SEQ ID NO: 6:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 341 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 6:

50

GATCAATGAT GCTACTATTG AAGCGGGCAG TTCCGTGTAC AGCTCCACCA AAGGCGATAC 60
 TGAATTGGGT GAAAATACCC GTATTATTGC TGAAAACGTA ACCGTATTAT CTAACGGTAG 120
 TATTGGCAGT GCTGCTGTAA TTGAGGCTAA AGACACTGCA CACATTGAAT CCGGCAAACC 180
 GCTTTCTTTA GAAACCTCGA CGTTGCCTC CAACATCCGT TTGAACAAACG GTAACATTAA 240
 AGGCGGAAAG CAGCTTGCTT TACTGGCAGA CGATAACATT ACTGCCAAA CTACCAATCT 300
 GAATACTCCC CGCAATCTGT ATGTTCATAC AGGTAAAGAT C 341

(2) INFORMATIONS POUR LA SEQ ID NO: 7:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 164 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SCUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 7:

GATCCAATG TTTGATTTTA CTGGCTGCTT CTCCATGCGC GGTATTGACC AAAGCCGCAA 60
 GGATATTGCG TTCCAGATTG TCTTCAGGC TGCCGCCGTT GACAGCGGTA TTAATCAGTG 20
 CGGCACTGCC CGCATTGGCT AGGTTGACGG TCAGGTTGTT GATC 164

(2) INFORMATIONS POUR LA SEQ ID NO: 8:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 219 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*
 (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 8:

GATCAATCAC ACATCTTGTGTC ATTTTTTCGA TTCCCTTCATT TCGGTTTCTA ATGTTTCAAT	60
TCTTGCAGGCC ATTTCCCTGAA TCGGCTTTAGT CAAAACGGGG ATGAACGCTT CGTATTGAC	120
GGTGTAGGTA TCCTTTGTTT TATTTACCAT CGGCAATCGA CCATATTCAT CTTCCAGCGC	180
AGCAATGTCC TGGGCAATAA ACCAATGCCG CAACCGATC	219

(2) INFORMATIONS POUR LA SEQ ID NO: 9:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 356 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 9:

GATCTTGGGT AAGCCCCCAA CCTGCATAGA AAGGCAGGCC GTAGCAGCTG ACTTTTTTGC	60
CGCGCAACAA GGCTTCAAAA CCGGTCAGCG AAGTCATGGT ATGTATTTCG TCTGCGTATT	120
GGAGACAGGT CAGGATGTGCG GCTTGTTCGG CGGTTGGTC GGCATATCGT GCAGGCATCAT	180
CAGGGGAAAT ATGGCCGATG CGGTTACCGC TGACTACATC GGGATGCGGT TTGTAGATGA	240
TATAGGCATT GGGTTTCTGT TCCCGTACGG TACGGAGCAA ATCCAGATTG CGGTAGATTT	300
GGGGCGAACCC GTAGCGGATA GACCGATCAT CTTCAACCTG GCCGGGAACG AGGATC	356

(2) INFORMATIONS POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 210 paires de bases

- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO. 10:

GATCCGCTTT CAGTTTCCGT ACCGGTGGCA TCAGTCAGT CCGTTTTGTG CACCAAACCG	60
CCTCCATATG AAACATAAAA CAAATCGCTT AAGCCCAAAG GGTTATCGAA CGATAAAGCG	120
ACATTTCCCTT GATATTGCC GGTCTGTTTG CCGCCCGCAT CATCTATACC GATACTGAAC	180
CGTATGGGTT TATTCTGCTG CCATTTGATC	210

(2) INFORMATIONS POUR LA SEQ ID NO: 11:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 259 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 11:

GATCCGAAA CGCAATTGGT CGAAAGCTAT ATGCTGAACG ATGTGTTGCG GTTTTGGGAC	60
AGCGCAGGTT TGGGCGATGG GAAAGAAGCC GACCGCGCCC ATCGCAAAA ACTGATTGAT	120
GTCCTGTCTA AACCTATAC TCATTCGGAT GGGCAGTGGG GCTGGATAGA TTTGGTGTTC	180
GTTATCCTTG ACGGCAGCTC CCGCGATTTG GGTACGGCCT ATGATTGTT GAGGGATGTT	240
ATCCTTAAAA TGATTGATC	259

(2) INFORMATIONS POUR LA SEQ ID NO: 12:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 436 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

GATCAAATGG ATGATTATA TAGAATTTC TTTACGACT GCGTGCCGTT TGAAAAGAAA	60
ATGCACAATC CCGTATCTCA TCGTGCCATA GATTTTCAA AGACTCCGGA AGCCATATTT	120
CGTTGCAATC TGCATACCGA ATTGAAGAAG AAGCGTAAAT TAGCGTTACG TTTAGGCAAG	180
CTGTCGGACA ATACAGCATG GATATTTAAA CCCCAGTCA TGAAAAATCT TCTGAAAAAC	240
CCGTCAACTC AAATTACCGA AAACGATGTC GTGCTCGATG TTAAACAAAA AGGTGTAGAT	300
ATGCGTATAG GCTTGGATAT TTCATCTATT ACCTAAAAAA ACAAGCCGA TAAAATCATC	360
TTGTTTCTG GTGATTCCGA TTTTGTCCCA GCAGCCAAAT TAGCCAGACG GGAAGGTATC	420
GATTTTATTC TTGATC	436

(2) INFORMATIONS POUR LA SEQ ID NO: 13:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 363 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 13:

GATCGTTTTA CGTCGCAATC GAGCTTGTG GTGCGCTGCG CTAAAAGCCA ATCTTCTCTC 60
 AATGGCCTGG GTGCCATTTC GCAGGGCACA GGTTTGCCTG GTGCGCAAGA CGATATTTAT 120
 ACCGTGCAGG AATATATGCA GTCGCGTTCG GCTTTGGATG CTTGCCTAA GAAAATGCC 180
 ATTGCGATT TTATGAAAAA AGAAGGCGAT ATTTTCAGCC GTTTTAATGG TTTTGGCCTG 240
 CGTGGCGAGG ATGAGGGT TTATCAATAC TACCGTGATA AGGTATCCAT CCATTTGAC 300
 TCTGTCTCAG GCATTTCCAA TTGAGCGTT ACATCGTTA ATGCCGGTGA ATCTAAAAG 360
 ATC
 363

(2) INFORMATIONS POUR LA SEQ ID NO: 14:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 314 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 14:

GATCTTGCCT CATTATATC TTCAACGATA TTGCAATTAC CGCCGTTCCA GTTGAATAA 60
 CAACGACTAA AATTGTAGTT CCTAAAAGAA TCATTCCTAT TCTTGCCTAC CATTTCCTAA 120
 TAATTGCGCC CGACAATTTC CATTAAATGC TCCATCAGTT CTTTACTTC CGGAAATCTG 180
 CTGTAATCTG ACATAAGACG CATAATTGAA CTATCAACGC CGTAACAGCC ATAGGTTTTA 240
 ATACCGTTTT CGGCCTGTTT CCAAATGCAA TTACTGTATT CGTAGCCTTT TACAAATTAA 300
 TCGGTTTCGG GATC
 314

(2) INFORMATIONS POUR LA SEQ ID NO: 15:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 256 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 15:

GATCATAACGA ATCTACCCCTA AAATACCCCG TCGCCGATTT AGGATTGGCT ACATAAAGCT	60
CATTATAAGG GTATTTTGAT GACATGATAAC GGTTAAATTTC ATTGCCGTTG TTTATCCTGA	120
TTCTATAAAAT TGGTTCAACA GCAAAGCCTC TGGATTCCCT TAATTGATTA TAATATTGCC	180
TGTATGTTTG TACATCATGT CTTGTCCACG GCTCTCCAGG AGTCCTCAGA ATAGCAATCC	240
CGTTAAATTTC CGGATC	256

(2) INFORMATIONS POUR LA SEQ ID NO: 16:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 235 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 16:

GATCCACGCC TGTGCCTACC TTGGCTTTTT GTTCGCCAAA CAAGGCATTT AAGGTTGAGG	60
ACTTGCCGAC ACCTGTCGCA CCGACAAGCA AGACATCCAA ATGACGGAAA CCGGCTGCTG	120

TGACTTTTG CCGGATTCA GAAATACGGT AACGATGCAT ATGCGCTCCT ACCAGCCAAA 180

AAAAGAAGCA ACCGTGCTAA TCGCCCCCTCC AATCGCTTT GCAGCACCGC CGATC 235

(2) INFORMATIONS POUR LA SEQ ID NO: 17:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 259 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 17:

GATCCAACGG GCATCGCTGT CCTTACTCGG TGTGGTTTGA CCGCTGATTT GTCCTTCCTC 60

GTCAACTTCT ATGGCCTGAC GCTGTTTGCT GCCGGCGGTC TGGATAATGG TGGCATCAAC 120

GACGGCGGCG GATGCTTCT CTATTTTCTAG GCCTTTTCG GTCAGTTGGC AGTTAACAG 180

TTTGAGTAAT TCGGACAGGG TGTGCTTGT CGCCAGCCAG TTGCGGTAGC GGCATAAGGT 240

ACTGTAATCG GGGATGATC 259

(2) INFORMATIONS POUR LA SEQ ID NO: 18:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 201 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE. SEQ ID NO: 18:

GATCTGTGCC GTTGATTTA TCTTCAGAT GCAGCATCGA ATATCGGAAA GCCAAATCAG	60
CAATTCTTT TGCATCGTGT GGATTTGAG ACGGGCCTAA TGACCGTACC CGCTTAATAA	120
AAAATGCACC GTCAATCAAA ATGGCGGTT TCATATTGCT TCCCCTATAT TTGTCAAAGA	180
TATAAAAAAG CCCTTGGGAT C	201

(2) INFORMATIONS POUR LA SEQ ID NO: 19:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 334 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 19:

AATTCAAAGG AGGCATTGT TGCAAGAAAA GTACAAAGTG ATTTGCAAAA AGCATTGAAT	60
GCTAGCAACT ATAACAAGCA GCAATATGCA AGACGTGCGG CAACAGCGTT AGAGAATGCT	120
TCAAAATCAA AAGTTATGGC AGCGAATTCT TTTGATCTA TCTTGTGCGA ACGGGTCAA	180
TATTCCTCGT ACATTGAGTT AATCGTACCA ATCGCCCTAA CCACATTTTC ATCAGAAAAT	240
ATGGAAATAA TAGCATCCCT ATACGCACCT AGTGTAAATAT TGTTTCTATT ATTAGTTATA	300
GCATTATTCTG AATACATAAT AGCACCTCCA AATT	334

(2) INFORMATIONS POUR LA SEQ ID NO: 20:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 238 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 20:

AATTCTGCG CACCTTGCC GATGGGGAGA TAATGCCCTT TTTGCAGCAT TCTGCCCTGA	60
TGGCCGCCGA AACCGGCTTT CAGGTGGTA CTTCTCGAAC CCATCACTTC CGGCACATCA	120
AATCCGCCCCG CCACGCACAC ATAGCCGTAC ATGCCCTGCA CGGCACGGCAC CAGTTCAAG	180
GTCTGCCCTT TGCAGGGCGGT ATAACGCCAA TACGAATAGA CCGGTTCGCC GTCCAATT	238

(2) INFORMATIONS POUR LA SEQ ID NO: 21:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 249 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 21:

AATTGGCGA GATGCTGCCG GAAACGGATT TAAAACAGAT TGCGGCGGCA GTGTTGAAGA	60
CGAACGATGA GGCGGCATTG CAGAAGGTGG TGAAAACGGC CAAAGGCAAT GCGCGGAAAC	120
TGTGAAAGCT GCTGCTGATT GTGGACTATT TGTTGCAGGT TAACCCCTGAT GTTGATTGG	180
ATGATGATGT AATCGAACAC GCGGAAACCT ATTTAATCCA CTAAACCTTT GACAGATAAG	240
GCAATAATT	249

(2) INFORMATIONS POUR LA SEQ ID NO: 22:

59

(1) CARACTERISTIQUES DE LA SEQUENCE

- (A) LONGUEUR: 212 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 22:

AATTTATGTA CGGTTTGCC GTTTGCAGTC AGCCAGTCGG CAAGGCGCAG AAAAAAAATCG	60
CCGACAGGGC CTTGAAGCAG CAGGATATTT TCTGGCTTT CAAGCAGGTT TTGCAGGTTA	120
TTTTGAGGA CGGTCTGTTT CATGTTGCAA TGTGGTTTG TTTTTATGT AATAGTTTA	180
GGTTGAACCTT TCAAGCATAc GCCAAGAGAA TT	212

(2) INFORMATIONS POUR LA SEQ ID NO: 23:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 227 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 23:

AATTCACTGC CTGCGTCATA TCACGGCTAC CTTGTGGTTC AGGGTTACTG TATCGCCCGC	60
GGCATCGACG GCTTCAATAT GCAGCTTCAG CCAGCCGTGC TGCGGGGGCGG ATGCGGTTAC	120
TTGGATGGAT TGGGCGCGTT TGGACTGAAT CACGGGCTGC AAGGCTTGCT CGGCGTACTG	180
TTTGGCCAGT ACTTCGATGC GCTTTAAATG CTTTGGCGG CGCAATT	227

(2) INFORMATIONS POUR LA SEQ ID NO: 24:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 167 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 24:

GATCCAGGAC TCAAAAACCG ATTTCCTAAT AGAGTGTCTA ATATCCAAT CTTTTTTACC	60
CCCTCTGCTG TAGAATTGAT AGAGAAAGTT TGTCTATCTT TTTCATATAC CCATGCCCTTC	120
TTTTTATCAT TGTAGCTAAC ATAACCGCCA AACAAATGCTT CTAGATC	167

(2) INFORMATIONS POUR LA SEQ ID NO: 25:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 251 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 25:

AATTCTTGCG GCCATTTCTT GAATGGCTTT AGTCAAAACG GGGATGAACG TTTCGTATTG	60
GACGGTGTAG GTATCGTTTG TTTTATTTAC CATCGGCAAT CGACCATATT CATCTTCCAG	120
CGCAGCAATG TCCCTGGCAA TAAACCAATG CGCGAACCGA TCTTCTTTAT GACTGCCGTC	180

61

CTTGATTGGA TTTCGCCACC ATTCGCGGAC TTTGTCCGCT CGTTCATCTG CCGGCAAGTC 240

TTTGAATAAT T 251

(2) INFORMATIONS POUR LA SEQ ID NO: 26:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 207 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 26:

AATTCCCGAC TATCGCGGAT GCGTAGTTTT TGCCGGTGGG CAAGAGCAGG TGTGGGATAA 60

GTTAGGTGAT TTGCCCGATG GCGTCAGCCT GACCCCGCCT GAATCGGTAA ATATTGACGG 120

CTTAAAATCC GTAAAACTCG TCGCATTAAA TGCTGCCGCT CAGGCTTTTA TTAACAAGCA 180

CGCCGGTATC GACAGCGTAC CTGAATT 207

(2) INFORMATIONS POUR LA SEQ ID NO: 27:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 379 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 27:

AATTGTTTGG GAATAATCCA AACAAACAGC ATCAGGATAG CGGCGGGCGCT CAGGCTGCCT 60

62

GAAAGGATTT TGCCCCGGTT TTTTGTAGGC AAAGCGGACG AGAAACCAAA GCAACAGCAG	120
CATGGTGTCC CAATAGCCGA TTGAGAATAG GATGGCCAAA CCTTCTAGGA AATGGCGTAA	130
ATCGTTTGTG GTAACCATGG GTAGTTCTG TGGTTAAATG TGCAGGCTGC TTTTGGCGA	240
ACCTTGCCGC ATCTCAAAAG CAGCCTGCGC TTCAGCGTTG CGTTACGCAG TAAAATAATG	300
AATAATTTGTA ACGGCTTGGG TATTTTTGT CAATATTCCC GCCTTCCCT TAACAGCTGC	360
CGCGCTTTCC GTTAAAATT	379

(2) INFORMATIONS POUR LA SEQ ID NO: 28:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 274 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADN (génomique)

- (vi) ORIGINE:
 - (A) ORGANISME: *Neisseria meningitidis*
 - (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 28:

AATTGCCGA AATCAGGCTG CTGCTCGATA ATCGGCCGG CCGATTGGCG TTGTGCCTCG	60
ATTAATCCA TCTTGTCTTG CAGACGTTTG GCCTGGCTT TGC GGCGGGCG TTGGCCAGT	120
TGTTCCATCC GCGTTTCCGC AAATGCCGCC CGTTTGTGCG CGTTGAATAAC CGCTTIGCAA	180
ATCACCTTGC CCTGCATATC CTTACAATC ACATGGTCGG CATCGTGGAT GTCGTAAGCC	240
ACCCGTACCT TCTGACCGCT GTAATCCAGC AATT	274

(2) INFORMATIONS POUR LA SEQ ID NO: 29:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 263 paires de bases
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple

63

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*

(B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 29:

AATTCCGTTTC TTATTGGGCT TTTTCCATCC ATCGGGTATG CCTGAAGGGA ACGCAAACCC	60
TGCCACTTGC CCATCGCTCC ATTCCCGCAT TAGCGCGTCT GACGGCAAGT GTTCTCGCGC	120
CCAATCAAGC CACGCCTGCC GCATTGCGGC CTTGCTCTGC TGAAAACCTTC GCAGTGCTTT	180
TGCAACCGGC CCATCATTAA CTTCAATCAA ATAAATCATT ATATTTGGT TCATTTTCC	240
TACACCTTCG CCACATCCAA ATT	263

(2) INFORMATIONS POUR LA SEQ ID NO: 30:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 316 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*

(B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 30:

AATTGTTCAA GAAAAAAAGTC GGCACGGCGC GGCAACGGGG AAAATGCGTT GACGCCGTCT	60
TTTCTAAGG TGATGTAGTA GGGGCGGAAA TAGCCTTCTT CAAACGCCA GAAACTGGCT	120
TGGTTTTCGT TTGCAATGCG TTTGCAATG ACGTGATAAG GGCGTGTGTC GCCAAAGCAG	180
ACAACGGCCT GGATGTGATG TTGAGTGATG TATTCTTGCA AAAACTCAGG AAAGGGCGTCG	240
TAGTTGTCGT TAAAAACAAAC GGTATGCGCT TGAGTGGGCG GATAAAAATA GTCGTCGCCT	300

GCATTAAGT TGAATT

316

(2) INFORMATIONS POUR LA SEQ ID NO: 31:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 324 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(iii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 31:

AATTCAATCA	ACGGAAAACA	CATCAGCATC	AAAAACAAACG	GTGGTAATGC	CGACTTAAAA	60
AACCTTAACG	TCCATGCCAA	AAGCGGGGCA	TTGAACATTC	ATTCCGACCG	GGCATTGAGC	120
ATAGAAAATA	CCAAGCTGGA	GTCTACCCAT	AATACGCATC	TTAATGCACA	ACACGAGCGG	180
GTAACGCTCA	ACCAAGTAGA	TGCTTACGCA	CACCGTCATC	TAAGCATTAC	CGGCAGCCAG	240
ATTTGGCAAA	ACGACAAACT	GCCTTCTGCC	AACAAGCTGG	TGGCTAACGG	TGTATTGGCA	300
CTCAATGCGC	GCTATTCCCA	AATT				324

(2) INFORMATIONS POUR LA SEQ ID NO: 32:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 230 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(iii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SOUCHE: Z2491

65

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 32:

AATTATGCAA AAAAACGCAA CGCCGAAAAA CTGGCACCGC GCGGATATTG TTGCTGCTTT	60
GAAAAAGAAA GCCTGGTCAC TTGAGGACT TTCAATAGAA GCGGGGTTGT CGCGAATAC	120
GCTTAGAAGC GCAGTGGCCG CCCCTTATCT TAAGGGAGAA AGGATTATTG CCGCTGCAAT	180
CGGAGTGGAA CGCGAAGAGA TTTGGTCCGA ACGGTATGCA GATCGGAATT	230

(2) INFORMATIONS POUR LA SEQ ID NO: 33:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 249 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

- (A) ORGANISME: *Neisseria meningitidis*
- (B) SCUCHE: 22491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 33:

AATTTAATCG GTGGAATGCC TGTTCAACCG CACCAATCCC GCTGAATACG GTTGCTAATC	60
TAATATGTGA ATCAGGTTTA AGAAAAGTTT TAGATTTCCA ACCTTGTGTA CTGGGAAAGA	120
GCAAAGTTTT TTGTAATCGA GTATCGTGTG TCTGTGCCAT TGTCGAAATA GTCATACTTA	180
TATCGTTCTG TTTATCTTAT CAATATGAAA ACTACATCGT TGATTGCCCT GACAATGCCT	240
TGGTCAATT	249

(2) INFORMATIONS POUR LA SEQ ID NO: 34:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 343 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

66

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*
 (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 34:

AATTCTTGTC CCGGAGTCCA ACGTATATTT ACCCTCCTGC GAGCTAAAAG ACTATTATTC	60
TCCACTGCCA CAGTAGCCGC ATTACCCGCC GTATTACAT CCCCTTTAAC CAATGCCACT	120
GCGCTGCCCTG CGATAATCTG CGAGTAGGCT ATGACTTTTT GGCCTTCTTG GGGTGACAGT	180
TTGCCTACAT CGCCTCCGTC CAACAGGGTT TCTCCACCA TCTCGCCGAC TGCCGCGCCG	240
ATTGCGCCGT CCCGACATTT GCCTTTATTT GCTACGGCG ATGCACAGCC TGCTACGGCA	300
TGGGCTATCT TGTGGGCAAT GTAGTCTTCG CTGAGATTAA ATT	343

(2) INFORMATIONS POUR LA SEQ ID NO: 35:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 184 paires de bases
 (B) TYPE: nucléotide
 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(vi) ORIGINE:

(A) ORGANISME: *Neisseria meningitidis*
 (B) SOUCHE: Z2491

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 35:

AATTCTTCAA ACATCGTTTC GATAATCGGG TCGGTGTACA CACTGATGCG GTCGCCCGCA	60
CGGCTTTGAC CGGCTCGGAA AATATAGGCG GTGGCTTTGC CGTCGGCGAT GTCGACCGAC	120
CAACGCCAGA TGGCGTCTTC GGTATTCAA CAATCACCCG CACAGCTTTC ACCTGCGCGG	180
AATT	184

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 36:

TATGCTCAAT CTCATTTC AATGCAAAA CTTTCTGAT TTTCTACT TTTGCTCAA	60
TATTAGGAAG GTTTAGGCA ATTGAAAATT TTTGGCGCA TTTTATGCG TCAAATTTCG	120
TTAACAGACT ATTTTGCAA AGGTCTCCST CTGAAAAGC AAGGATAGGG CATCTGCCCT	180
TTTGATTGTT TGATTAACGA TACAAGGAGT TTCAAAATGA GAGTTTATA GTGGATTAAC	240
AAAAACCACT ACAGCGTTGC CTCGCCTTGC CGTACTATTT GTACTGTCTG CGGCTTCGTC	300
GCCTTGTCT GATTTAAATT TAATCCACTA TATGTGTCA TGAAATGACT TGGTCTGGAG	360
GCTCAGGTAA TCGCAACAA AGTCATATT ATTGCGAAAT TTGCGAATCT GCAGGGCTTA	420
ACGATACGGG AAATCCTGAT AAATCTTAG GATTGCCAAA CAATACGTT AGTAATCCGC	480
CTCGTTGGGG AGCTACAATC GGAGCTTAG CAGGTAGCCG CATAAGTATG CCTGAATTG	540
GTACGTTTGC GAGCCATGCC ATTGAAAATT TCGACTGGTC ATGGTATCGA CGTTATAGGG	600
AAATTGCCGA AACGATTGAA CGAGAATATT CAGGCGGTTT GCCTTAATAG TTGAGGAGGT	660
CATGATGTTT GCCAAACATT ATCAATTCA CGCACTCGGC ATCATGCTGC TTCTTATAT	720
GTGATTCTC TATACGACCG ATTTCCTCAA TCTGACGTAT TGGATGCTGT TTTTATCTG	780
TTTATTACA GGAAAATAT TAGCTGTTT GTTAGAGAAA AGCTTTAAAT AAAATAGCAG	840
CTAGTCGCAA AAGTCGTCT GAAACCTTTT CAGGCGGCCT TTCTAAAATA CATCCAACCT	900
CCTAATCCCT ATTTCCTCAA AAGGAAATCT ATGCCCCATC TGCAAAACCT GTCTTGGC	960
TTAAAGAAAA AGCTGCCTGT TATCCTGCAA ACAGAAATAT CAGAATGCGG CTTGGCATGT	1020
CTGGCGGCTG TGGCGGGATT TCATGGTTTC CATACTGATT TACGCGCACT GCGTCAAAA	1080
TACTGTCCGA GACCTTGCA AAATTCCCCA AAATCCCCA AATGTCTTGG TGGGAATT	1140
GGGGAAATTG GCAAAGGTCT CATTCTATAA CTGAAATAC TTTAAATTG ATGACAAAAT	1200
AGTAAATATT GCTAAAATAA TATTGATGTC ATGAAATTG TTCTGCTCC ATGCTGTTG	1260
GTATCCTGG CTGTCATACC CCTTAAACC TTAGCTGCCG ATGAAAACGA TGCAGAACTT	1320
ATCCGTTCCA TGCAGCGTCA GCAGCACATA GATGCTGAAT TGTTAACTGA TGCAAAATGTC	1380

CGTTTCGAGC AACCATTGGA GAAGAACAAAT TATGTCCTGA GTGAAGATGA AACACCGTGT	1440
ACTCGGGTAA ATTACATTAG TTAGATGAT AAGACGGCGC GCAAATTTTC TTTCTTCCT	1500
TCTGTGCTCA TGAAAGAAC AGCTTTAAA ACTGGATGT GTTAGGTTTC CAATAATTG	1560
AGCAGGCTAC AAAAAGCCGC GCAACAGATA CTGATTGTGC GTGGCTACCT CACTTCCAA	1620
GCTATTATCC AACCACAGAA TATGGATTCCG GGAATTCTGA AATTACGGGT ATCAGCAGGC	1680
GAAATAGGGG ATATCCGCTA TGAAGAAAAA CGGGATGGGA AGTCTGCCGA GGGCAGTATT	1740
AGTGCATTCA ATAACAAATT TCCCTTATAT AGGAACAAAAA TTCTCAATCT TCGCGATGTA	1800
GAGCAGGGCT TGGAAAACCT GCGTCGTTG CCGAGTGTAA AAACAGATAT TCAGATTATA	1860
CCGTCCGAAG AAGAAGGCAA AAGCGATTAA CAGATCAAAT GGCAGCAGAA TAAACCCATA	1920
CGGTTCAAGTA TCGGTATAGA TGATGCGGGC GGCAAAACGA CCGGCAAATA TCAAGGAAAT	1980
GTGCGTTTAT CGTTCGATAAA CCCTTTGGGC TTAAGCGATT TGTGTTATGT TTCATATGGA	2040
CGCGGTTTGG TGCACAAAAC GGACTTGACT GATGCCACCG GTACGGAAAC TGAAAGCGGA	2100
TCCAGAAGTT ACAGCGTGCA TTATTCGGTG CCCCTAAAAA AATGGCTGTT TTCTTTAAT	2160
CACAAATGGAC ATCGTTACCA CGAAGCAACC GAAGGCTATT CCGTCAATTAA CGATTACAAC	2220
GGCAAAACAAT ATCAGAGCAG CCTGGCGGCC GAGCGCATGC TTTGGCGTAA CAGGTTTCAT	2280
AAAACCTCAG TCGGAATGAA ATTATGGACA CGCCAAACCT ATAAATACAT CGACGATGCC	2340
GAAATCGAAG TGCAACGCCG CCGCTCTGCA GGCTGGGAAG CCGAATTGCG CCACCGTGCT	2400
TACCTCAACC TTGGCAGCT TGACGGCAAG TTGTCTTACA AACGCGGGAC CGGCATGCC	2460
CAAAGTATGC CCGCACCTGA AGAAAACGGC GGCGGTACTA TTCCAGGCAC ATCCCGTATG	2520
AAAATCATAA CCGCCGGATT GGATGCAGCG GCCCCGTTTA TGTGGGCAA ACAGCAGTTT	2580
TTCTACGCAA CCGCCATTCA AGCTCAATGG AACAAAACGC CTTTGGTTGC CCAAGACAAG	2640
TTGTCTATCG GCAGCCGCTA CACCGTTCGC GGATTTGATG GGGAGCAGAG TCTTTTCGGA	2700

GAGCGAGGTT TCTACTGGCA GAATACTTTA ACTTGGTATT TTCACTCCGAA CCATCAGTTC	3760
TATCTCGGTG CGGACTATGG CCGCGTATCT GGCGAAAGTG CACAATATGT ATCGGGCAAG	3820
CAGCTGATGG GTGCAGTCGT CGGCTTCAGA GGAGGGCATA AAGTAGGCAG TATGTTTGCT	3880
TATGATCTGT TTGCCGGCAA GCGCTTCAT AAACCCAAAG GCTTCAGAC GACCAACACC	3940
GTTCACGGCT TCAACTTGAA TTACAGTTTC TAACCTCTGA ATTTTTTAC TGATATTTAG	3000
ACGGTCTTTC CTATCCTCA GACTGTCAA CTTACCTAC GTACTTGGCG CGCAGTACGT	3060
TCATCTTCAA AATGGAATAG ACATGAATAA AGGTTTACAT CGCATTATCT TTAGTAAAAA	3120
GCACAGCACC ATGGTTGCAG TAGCCGAAAC TGCCAACAGC CAGGGCAAAG GTAAACAGGC	3180
AGGCAGTTCG GTTCTGTTT CACTGAAAAC TTCAAGGCGAC CTTTGGGCA AACTCAAAAC	3240
CACCCTTAAA ACCTTGGTCT GCTCTTTGGT TTCCCTGAGT ATGGTATTGC CTGCCCATGC	3300
CCAAATTACC ACCGACAAAT CAGCACCTAA AAACCAGCAG GTCGTTATCC TTAAAACCAA	3360
CACTGGTGCC CCCTTGGTGA ATATCCAAAC TCCGAATGGA CGCGGATTGA GCCACAACCG	3420
CTATACGCAG TTTGATGTTG ACAACAAAGG GGCAGTGTAA AACAAACGACC GTAACAATAA	3480
TCCGTTCTG GTCAAAGGCA GTGCGCAATT GATTTGAAC GAGGTACGCG GTACGGCTAG	3540
CAAACCTAAC GGCACTGTTA CCGTAGGGCG TCAAAAGGCC GACGTGATTA TTGCCAACCC	3600
CAACGGCATT ACCGTTAATG GCGGCGGCTT TAAAAATGTC GGTCGGGGCA TCTTAACAT	3660
CGGTGCGCCC CAAATCGGCA AAGACGGTGC ACTGACAGGA TTTGATGTGC GTCAAGGCAC	3720
ATTGACCGTA GGAGCAGCAG GTTGGAAATGA TAAAGGCGGA GCCGACTACA CCGGGTACT	3780
TGCTCGTGCA GTTGCCTTGC AGGGGAAATT ACAGGGTAAA AACCTGGCGG TTTCTACCGG	3840
TCCTCAGAAA GTAGATTACG CCAGCGGCGA AATCACTGCA GGTACGGCAG CGGGTACGAA	3900
ACCGACTATT GCCCTTGATA CTGCCGCACT GGGCGGTATG TACGCCGACA GCATCACACT	3960
GATTGCCAAT GAAAAAGGCG TAGGCGTCAA AAATGCCGGC AACTCGAAG CGGCCAAGCA	4020
ATTGATTGTG ACTTCGTCAG GCCCGATTGA AAACAGCGGC CGCATGCCA CCACTGCCA	4080

CGGCACCGAA	GCTTCACCGA	CTTATCTCTC	CATCGAAACC	ACCGAAAAAG	GAGCGGCAGG	4140
CACATTITATC	TCCAATGGTG	GTCCGGATCGA	GAGCAAAGGC	TTATTGGTTA	TTGAGACGGG	4200
AGAAGATATC	AGCTTGGTA	ACGGAGCCGT	GGTGCAGAAT	AACGGCAGTC	GCCCAGCTAC	4260
CACGGTATTA	AATGCTGGTC	ATAATTGGT	GATTGAGAGT	AAAACTAATG	TGAACAATGC	4320
CAAAGGCTCG	GCTAATCTGT	CGGCCGGCGG	TCGTACTACG	ATCAATGATG	CTACTATTCA	4380
AGCGGGCAGT	TCCGTGTACA	GCTCCACCAA	AGGCGATACT	GAATTGGGTG	AAAATACCGS	4440
TATTATTGCT	GAAAACGTAA	CGTATTATC	TAACGGTAGT	ATTGGCAGTG	CTGCTGTAAT	4500
TGAGGCTAAA	GACACTGCAC	ACATTGAATC	GGGCAAACCG	CTTTCTTTAG	AAACCTCGAC	4560
CSTTGCCTCC	AACATCCGTT	TGAACAACGG	TAACATTTAA	GGCGGAAAGC	AGCTTGCTTT	4620
ACTGGCAGAC	GATAACATTA	CTGCCAAAC	TACCAATCTG	AATACTCCG	GCAATCTGTA	4680
TGTTCATACA	GGTAAAGATC	TGAATTGAA	TGTTGATAAA	GATTGTCTG	CCGCCAGCAT	4740
CCATTGAAA	TCGGATAACG	CTGCCATAT	TACCGGCACC	AGTAAAACCC	TCACTGCCTC	4800
AAAAGACATG	GGTGTGGAGG	CAGGCTTGCT	GAATGTTACC	AATACCAATC	TGCGTACCAA	4860
CTCGGGTAAT	CTGCACATTC	AGGCAGCCAA	AGGCAATATT	CAGCTTCGCA	ATACCAAGCT	4920
GAACGCAGCC	AAGGCTCTG	AAACCACCGC	ATTGCAGGGC	AATATCGTTT	CAGACGGCCT	4980
TCATGCTGTT	TCTGCAGACG	GTCATGTATC	CTTATTGGCC	AACGGTAATG	CCGACTTTAC	5040
CGGTACAAT	ACCTGACAG	CCAAGGCCGA	TGTCAATGCA	GGATCGGTTG	GTAAAGGCCG	5100
TCTGAAAGCA	GACAATACCA	ATATCACTTC	ATCTTCAGGA	GATATTACGT	TGGTTGCCGG	5160
CAACGGTATT	CAGCTTGGTG	ACGGAAAACA	ACGCAATTC	ATCAACGGAA	AACACATCAG	5220
CATCAAAAAC	AACGGTGGTA	ATGCCGACTT	AAAAAACCTT	AACGTCCATG	CCAAAAGCGG	5280
GGCATTGAAC	ATTCACTCCG	ACCGGGCATT	GAGCATAGAA	AATACCAAGC	TGGAGTCTAC	5340
CCATAATACT	CATCTTAATG	CACAACACGA	GCAGGTAAACG	CTCAACCAAG	TAGATGCCTA	5400

CGCACACCGT CATCTAAGCA TTACCGGCAG CCAGATTTGG CAAAACGACA AACTGCCTTC	5460
TGCCAACAAAG CTGGTGGCTA ACGGTGTATT GGCACCTCAAT GCGCGCTATT CCCAAATTGC	5520
CGACAACACC ACGCTGAGAG CGGGTGCAAT CAACCTTAACG GCCGGTACCG CCTAGTCAA	5580
GGCGGGCAAC ATCAATTGGA GTACCGTTTC GACCAAGACT TTGGAAGATA ATGCCGAATT	5640
AAAACCATTG GCCGGACGGC TGAATATTGA AGCAGGTAGC GGCACATTAA CCATCGAAC	5700
TGCCAACCGC ATCAGTGCAG ATACCGACCT GAGCATCAA ACAGGGGAA AATTGCTGTT	5760
GTCTGCAAAA GGAGGAAATG CAGGTGCGCC TAGTGCTCAA GTTTCCTCAT TGGAAGCAA	5820
AGGCAATATC CGTCTGGTTA CAGGAGAAAC AGATTAAGA CGTTCTAAAA TTACAGCCGG	5880
TAAAAAATTG GTTGTGCCA CCACCAAAGG CAAGTTGAAT ATCGAAGCCG TAAACAACTC	5940
ATTCAGCAAT TATTTTCCTA CACAAAAAGC GGCTGAACTC AACCAAAAAT CCAAAGAATT	6000
GGAACAGCAG ATTGCGCAGT TGAAAAAAAG CTCGCCTAAA AGCAAGCTGA TTCCAACCT	6060
GCAAGAAGAA CGCGACCGTC TCGCTTCTA TATTCAAGCC ATCAACAAGG AAGTTAAAGG	6120
TAAAAAAACCC AAAGGCAAAG AATACTGCA AGCCAAGCTT TCTGCACAAA ATATTGACTT	6180
GATTTCGGCA CAAGGCATCG AAATCAGCGG TTCCGATATT ACCGCTTCCA AAAAAACTGAA	6240
CCTTCACGCC GCAGGGTAT TGCCAAAGGC AGCAGATTCA GAGGCGGCTG CTATTCTGAT	6300
TGACGGCATA ACCGACCAAT ATGAAATTGG CAAGCCCACC TACAAGAGTC ACTACGACAA	6360
AGCTGCTCTG AACAAAGCCTT CACGTTGAC CGGACGTACG GGGSTAAGTA TTCAATGCAGC	6420
TGCGGCACTC GATGATGCAC GTATTATTAT CGGTGCATCC GAAATCAAAG CTCCCTCAGG	6480
CAGCATAGAC ATCAAAGCCC ATAGTGATAT TGTACTGGAG GCTGGACAAA ACGATGCCTA	6540
TACCTTCTTA AAAACCAAAG GTAAAAGCGG CAAAATCATC AGAAAAAACCA AGTTTACCA	6600
CACCCGCGAC CACCTGATTA TGCCAGCCCC CGTCGAGCTG ACCGCCAACG GTATCACGCT	6660
TCAGGCAGGC GGCAACATCG AAGCTAATAC CACCCGCTTC AATGCCCTG CAGGTAAAGT	6720
TACCCCTGGTT GCGGGTGAAG AGCTGCAACT GCTGGCAGAA GAAGGCATCC ACAAGCACGA	6780

GTTGGATGTC CAAAAAAGCC GCCGCTTAT CGGCATCAAG CTAGGTAAGA GCAATTACAG	6840
TAAAAACGAA CTGAACGAAA CCAAATTGCC TGTCCCGTC GTCGCCAAA CTGCAGCCAC	6900
CCGTTCAGGC TGGGATACCG TGCTCGAAGG TACCGAATTG AAAACCACGC TGGCCGGTGC	6960
CGACATTCAAG CCAGGTGTAG GCGAAAAAGC CCGTGTGAT GCGAAAATTA TCCTCAAAGG	7020
CATTGTGAAC CGTATCCAGT CGGAAGAAAA ATTAGAAACC AACTCAACCG TATGGCAGAA	7080
ACAGGGCGGA CGCGGCAGCA CTATCGAAAC GCTAAAATG CCCAGCTTCG AAAGCCCTAC	7140
TCCGCCAAA TTGTCCGCAC CCGGCGGCTA TATCGTCGAC ATTCCGAAAG GCAATCTGAA	7200
AACCGAAATC GAAAAGCTGT CCAAACAGCC CGAGTATGCC TATCTGAAAC AGCTCCAAGT	7260
AGCGAAAAAC ATCAACTGGA ATCAGGTGCA GCTTGCTTAC GACAGATGGG ACTACAAACA	7320
GGAGGGCTTA ACCGAAGCAG GTGCGGCGAT TATCGCACTG GCCGTTACCG TGGTCACCTC	7380
AGGCGCAGGA ACCGGAGCCG TATTGGGATT AAACGGTGCG GCCGCCGCCG CAACCGATGC	7440
AGCATTGCC TCTTGGCCA GCCAGGCTTC CGTATCGTTC ATCAACAACA AAGGGATGT	7500
CGGCAAAACC CTGAAAGAGC TGGGCAGAAG CAGCACGGTG AAAATCTGG TGGTTGCCGC	7560
CGCTACCGCA GGCGTAGCCG ACAAAATCGG CGCTTCGGCA CTGAACAATG TCAGCGATAA	7620
GCAGTGGATC AACAAACCTGA CCGTCAACCT AGCCAATGCC GGCAGTGCCG CACTGATTAA	7680
TACCGCTGTC AACGGCGGCA GCCTGAAAGA CAATCTGGAA GCGAATATCC TTGCGGCTTT	7740
GGTCAATACC GCGCATGGAG AAGCAGCCAG TAAAATCAAA CAGTTGGATC AGCACTACAT	7800
AGTCCACAAG ATTGCCCATG CCATAGCGGG CTGTGGCGA GCGGCCGCCA ATAAGGGCAA	7860
GTGTCAAGGAT GGTGGATAG GTGCGGCTGT GGGCGAGATA GTCGGGGAGG CTTTGACAAA	7920
CGGCAAAAT CCTGACACTT TGACAGCTAA AGAACCGAA CAGATTTGG CATAACGCAA	7980
ACTGGTIGCC GGTACGGTAA GCGGTGTGGT CGGCAGCGAT GTAAATGCC CGGCGAATGC	8040
GGCTGAGGTA GCGGTGAAAA ATAATCAGCT TAGCGACAAA GAGGGTAGAG AATTTGATAA	8100

CGAAATGACT GCATGGCCA AACAGAATAA TCCTCAACTG TGCAGAAAAA ATACTGTAAA	8160
AAAGTATCAA AATGTTGCTG ATAAAAGACT TGCTGCTTCG ATTGCAATAT GTACGGATAT	8220
ATCCCGTAGT ACTGAATGTA GAACAATCAG AAAACAACAT TTGATCGATA GTAGAAGCCT	8280
TCATTCACTCT TGGGAAGCAG GTCTAATTGG TAAAGATGAT GAATGGTATA AATTATTCAG	8340
CAAATCTTAC ACCCAAGCAG ATTTGGCTT ACAGTCTTAT CATTGAATA CTGCTGCTAA	8400
ATCTTGGCTT CAATCGGGCA ATACAAAGCC TTTATCCGAA TGGATGTCCG ACCAAGGTTA	8460
TACACTTATT TCAGGAGTTA ATCCTAGATT CATTCCAATA CCAAGAGGGT TTGTAACACAA	8520
AAATACACCT ATTACTAATG TCAAATACCC GGAAGGCATC AGTTTCGATA CAAACCTAAA	8580
AAGACATCTG GCAAATGCTG ATGGTTTTAG TCAAGAACAG GGCATTAAG GAGCCCATAA	8640
CCGCACCAAT TTTATGGCAG AACTAAATTAC ACGAGGAGGA CGCGTAAAAT CTGAAACCCA	8700
AACTGATATT GAAGGCATTA CCCGAATTAA ATATGAGATT CCTACACTAG ACAGGACAGG	8760
TAAACCTGAT GGTGGATTTA AGGAAATTTC AAGTATAAAA ACTGTTTATA ATCCTAAAAAA	8820
ATTTTCTGAT GATAAAATAC TTCAAATGGC TCAAATGCT GCTTCACAAG GATATTCAAA	8880
AGCCTCTAAA ATTGCTAAA ATGAAAGAAC TAAATCAATA TCGGAAAGAA AAAATGTCAT	8940
TCAATTCTCA GAAACCTTTG ACGGAATCAA ATTTAGATCA TATTTGATG TAAATACAGG	9000
AAGAATTACA AACATTCAAC CAGAATAATT TAAAGGAAAA ATTATGAAAA ATAATATTTT	9060
TCTAAACTTA AATAAAAAAT CTATAAATAA CAACCATTTC GTTATTTCGA TTTTTTTGAA	9120
AACAATTTAC CAATTGAAA CAAAGATAC GCTTTAGAG TGTTTTAAAA ATATTACAAC	9180
TACCGGACAT TTTGGAGTAA TAGGTGCTCA ATATGAAAAA ATAGATGCTA CCAGATGGAT	9240
TGGAGATTAT GAAGAGGTAA ATGGATTGTA GTATATTGAT AAAGCTCCTT CTATTTATT	9300
TTCAAGTTGGA GATGATTICA ATCCTGAAGA ATTAATTATA CCTATTAAATT TAGCATATCA	9360
TTACTTTAAT ATTGCAATAT CTGATTTCCTT AATAGCTCAC CCTGAATATC AAAAAAAAGTG	9420
TAAAGAAATA CAAAAAACAT ATTCTCAAAC AAACGTAGC CTGCATGAAA CCTAAAATCC	9480

ATCGCTAAGG TGTGTGCTTC AGCACCGCACG CGTTCCATGA TTTACGGCTC AATGCCGTCT	9540
GAAAAGCTCA CAAATTTTCA GACGGCATTG TTATGCAAG TAAATATTCA GATTCCCTAT	9600
ATACTGCCCA GACCGTGCAG TGCTGAAGAC ACCCCCTACG CTTGCTGCAG AACCTTCGGG	9660
TAAAACCGGT GTGAGCATTG GCGCACCGTA TGCCAATGAG AACAGTCGCA TCGTGTCAAG	9720
CACCA CGGAT ATCAGTTCGG AAAACGGCAA AATCAAATT CAATCTTACG GTGACCAATA	9780
TTACTATGCG AGACAGAGCG AACTCTATAC CTTTGAACGC CGCAGCTACA AAACCTGGCAA	9840
ATGGTACAAC CGCAAACACA TTACCGAAGT CAAAGAACAC AAAAACGCCA AGCCCCACGC	9900
AGTAAACCTC AGCGCATCCC AAGGCATCGA CATCAAATCT GGTGGCAGCA TCGACCCCTA	9960
CGCCACCGCA TTGATGCC CCAAAGGCAG CATTAAACATC GAAGCCGGGC GGAAATTGAC	10020
ACTCTATGCC GTAGAAGAGC TCAACTACGA CAAACTAGAC AGCCAAAAAA GGCGCAGATT	10080
TCTCGGCATC AGCTACAGCA AAGCACACGA CACCACCAAC CAAGTCATGA AAACCGCGCT	10140
GCCCTCAAGG GTAGTTGCAG AATCAGCCAA CCTCCAATCG GGCTGGGATA CCAAACGTCA	10200
AGGCACACAG TTTGAAACCA CACTGGTGG CGCAACCATA CGGGCAGGGCG TAGGTGAGCA	10260
GGCACGGGCA GATGCCAAGA TTATCCTCGA AGGGATCAAAG AGCAGCATCC ACACAGAAAC	10320
CGTGAGCAGC AGCAAATCTA CTCTATGGCA AAAACAGGCA GGACGGGGCA GTAACATCGA	10380
AACCTTGCAA TTGCCGAGIT TCACCGGTCC CGTTGGCCCC GTACTGTCCG CACCCGGCGG	10440
TTACATTGTC GACATTCCGA AAGGCAATCT GAAAACCCAA ATCGAAACCC TCACCAAGCA	10500
GCCCGAGTAT GCTTATTGTA AACAACTTCA AGTTGCGAAA AACATCAACT GGAATCAGGT	10560
GCAGCTTGCT TACGATAAAAT GGGACTACAA ACAGGAGGGC ATGACACCCCG CAGCAGCAGC	10620
TGTCGTGTT ATCGTCGTA CCGTATTGAC CTACGGTGCA CTGTCGCCCC CGGCAGCCGC	10680
CGGAACGGCG GGCGCGGCAG GCGCAGGAGC GGGAGGAGCC GCAGCAGGAA CGGCAGCCCG	10740
AACTGGAGTA GCAGCAGGAA CGGCAGCCAC AACCGGAGTA GCAGCAGGCA CATCAGCTGC	10800

75

AGCTATCACC ACAGCCGCAG GCAAAGCCGC ACTGCCAGT CTCGCCAGCC AAGCCGCAGT 10860
 TTCCCTCATC AACAAACAAAG GAGACATAAA CCATACCCCTG AAAGAACTGG GCAAAGCAG 10920
 CACCGTCAGA CAGGCCGCAG CCGCCGCCGT AACCCGAGGC GTACTGCAGG GCATAAGCGG 10980
 GCTGAACACC CAAGCAGCCG AAGCCGTCAG CAAACATTTT CACAGTCCCG CAGCAGGCAA 11040
 ACTGACCGCT AACCTGATCA ACAGCACCGC TGCCSCAAGT GTCCATACCG CCATCAACGG 11100
 CGGCAGCCTG AAAGACAACT TGGGCGATGC CGCACTGGGT GCGATAGTCA GTACCGTACA 11160
 CGGAGAAGTA GCGAGCAAAA TCAAATTTAA TCTCAGCGAA GACTACATTG CCCACAAGAT 11220
 AGCCCATGCC GTAGCAGGCT GTGCATCGGC GGTAGCAAAT AAAGGCAAAT GTCCGGACGG 11280
 CGCAATCGGC CCGGCAGTCG GCGAGATGCT GGGAGAAACC CTGTTGGACG GACCGATGT 11340
 AGGCAAACGT TCACCCCAAG AACGCCAAAAGTCA TAGCCATGCC TACTCGCAGA TTATCGCAGG 11400
 CAGCGCAGTG GCATTGGTTA AAGGGGATGT GAATACGGCG GTGAATGCCG CTACTGTGGC 11460
 AGTGGAGAAT AATACTCTTT TAGCTCGCAG GAGGGTAAAT ATACGTTGGA CTCCCGACAGA 11520
 AGAATTGGAA CATGAATATG CCATTCTGA AATCCAGGCC ATTACCAATC AAATCCGAAG 11580
 GCTGGATCCG AAATTTAACG GGATTGCTAT TCTGAGGACT CCTGGAGAGC CGTGGACAAG 11640
 ACATGATGTA CAAACATACA GGCAATATTA TAATCAATTAGGAAATCCA GAGGCTTTGC 11700
 TGTTGAACCA ATTTATAGAA TCAGGATAAA CAACGGCAAT GAATTTAACCGTATGTC 11760
 ATCAAAATAC CCTTATAATG AGCTTTATGT AGCCAATCCT AAATCGGCCA CGGGGTATTT 11820
 TAGGGTAGAT TCGTATGATC CTGCGACAAG GGAAATTATT TCAAGAAAAT TTACCCATT 11880
 TTCTCAAATC CAAGAAAGTA CGGGGATTGG TTATATCAAG GAGGCTGTTA GAAAATATAG 11940
 CCCTGGTACT GTCATTTCCA ATGTTCCAAG TACACCTACT ACGATAAGAG GAAGAAAGCT 12000
 TGAAGGAAAAA CTTATTTAG AAGTTCTGC TCAGGTCAAT CCAATTCCAC AATCTGTATT 12060
 AAGGGCGGCA CAAGAAGAAA ATGTTATCAT TAGAGATACA ACAGGAAGGA TTTACAAATG 12120
 AAGAAAGATA TTTTATTTG TGAGCAGTGG TCTTATGGTT ATAAGAGACT TCATAAGCCT 12180

TTTTCTGAGA AACAAGCTGA GGAAAAACAT CTTAAAGGGG AGTTATATAC TGCCGTAATA	12240
GGTCGGCGA CACAACCTGA ATATGTAATT ACCTTGCAG AGGAAGTAGG TTTTTTTTCTG	12300
GTAAATTTTT TCGATAAAATT TGGAAGGGAT TATTTAACCC ATCAATTTCATTTCA AAAATATTCC	12360
AATTCGAATT ATTATTTCT TTCTATGGCT GTATGGAGAG ATTATATAAC TTTGGAATCT	12420
CATGACTTAG CAGAAGGATA TACTTATTTCT TTCAATGAAA ATACGGATGA TTGCTATGTT	12480
TTGAAACAAAG ATTTTATTAA TAATGAGCGA TATGAAAAAA CAGAATTATA TTCCCAAAAAA	12540
GATAAGGTAATTTCTTCC AAAGTTGGT GAATATGATT TGGTGTAAA TCCGGACATT	12600
ATTTAATTAA GTTTTAAGGC CGTCTGAAAA AAATTTCAAA CGGCTTTTAT TATTGGGTTT	12660
GGAATCTGAG GATAAAGCTG ATAAAAACCA GGAAATTATC AGATTGCTAT ATACGTATTG	12720
TTGTACAGAC TAAAGGCAGC AATCAAATCA CTATTGCTTA CCCACAAAAA TAAATTGATT	12780
ATATGGAATA ATCATGAATA AGAGAAATGAA AATGTGTCTT GCTTGTCAAC AAGGCTATCT	12840
CTACCATTCG AAACCTAAAT ATCTTCATGA TGAAATTATT CTGTGTGATG AATGCGATGC	12900
AGTATGGCTC AAAGGTATGA ATATATTITA TGGAGAATAT GAAAAAGATT TTTATTCTTA	12960
TGTTCCCTTC ATGGAATCCC AAGGTATAAC GAGTGAATGT ATTTGGGAAG GAGATTGTT	13020
TGATCATCCA TATTATGAAG ATGAAAACTC AAATGATATG GATTGATGGA AATTTTAAGC	13080
CTGCGTAGGT ACGATTAGCC ATCAAACGGC GTAATCATAAC GCAAGATTAT CAACAGAGAG	13140
GGCTGGCAGC GATATACCAAC CCACAAGATT GCCCATGCCA TAGCGGGCTG TGCAGCAGCG	13200
GGGGCGAATA AGGGCAAGTG TCAGGGATGGT GCGATAGGCG CTGCAGTCGG TGAGATTGTT	13260
GGTGAGGCTT TGGTTAAGAA TACTGATTTC AGTCGTATGA GTGCGACCGA AATCGAAAAA	13320
GCTAAAGCGA AGATTACTGC CTATTCAAAA CTGGTTGCCG GCACTGCGTC TGCCGTGTA	13380
GGCGGGGATG TGAATACAGC GGCGAATGCG GCACAGATAG CGGTGGAGAA TAATACTTTC	13440
TATCCTAGAT GCGTTGGTGC AAAGTGTGAT GAATTTCAAA AGGAACAACA AAAATGGATA	13500

77

CGTGAAAATC CTGAAGAATA TCGAGAAGTT TTGCTTTTC AGACAGGATT TATTCCAATT 13560
 ATCGGTGATA TACAGAGTTT TGTACAAGCA CAGACCGCTG CCGATCACCT GTTTCCTTTG 13620
 CTGGGTGTGG TTCCGGGTAT CGGTGAATCG ATACAGGCCT ATAAAGTAGC GAAAGCGGCA 13680
 AAAAATTTAC AAGGCATGAA AAAAGCCTTG GACAAGGCAG CAACCGTTGC CACTGCACAG 13740
 GGCTATGTCA GCAAAACCAA AATCAAATC GGTCAAACGT AATTAAGGGT TACTGCAGCA 13800
 ACTGACAAAC AATTGCTGAA AGCTATTGGC GAAGGAAGGG ACACGACAGG TAAAATGACC 13860
 GAGCAGTTAT TTGACTCTTT AGCTAAACAA AATGGCTTCA GAGTGCTTTC GGGCGGCAAA 13920
 TACGGCGGAA ATAACGGTTT TGATCATGTA TGGCAGGCTG CCGATGGTAG TGTCGTTTG 13980
 ATTGTAGAAA GTAAGCAGAT TAGGAAACGGT ACGGTACAGC TGAATCGAA TGGTGCSSGT 14040
 GGATATACGC AAATGAGTGA GGATTGGATT AGACAAGTT TAGATCAATT ACCCGATGGT 14100
 AGTCCCGCTA AAGCTGCTGT CTTCAAAGCA AATAAGAACG GCACATTAAG AACAGCAATA 14160
 GCAGGGCTTG ATCGTCAAAC AGGTAAAGGCC GTTATTCTTC CTGTCAAAGT TCCTTCTAAA 14220
 ACCAATATAA GGAGATAACA ATGGGGCACA ATATGATGAC CACCCAAAAA TGGTATGAGC 14280
 ATATTACTAA TGTAATCATA GGCAATACTG CTAATTCAA TAGCGTTGC CTTGACTCTA 14340
 TAGATTATGT AGATGAAAGA AAAGGCCTTC CGCTTGCAGC TATGCAACAT ATTTCATGG 14400
 ACGTTAGAGC TGCAGCTTCC CATGCCTATC TATTTGAACA TGATCTTAAG AAATTCAAGC 14460
 AATATGCTTA TGTTGCAGGA AAGCTGGGGG TTTTGCTGAG TGTAAATTCT ACAGACCCCTG 14520
 AACCCCTCTT CTTCCCTGT GACATGCTCA ACATTCAAAA TCCGATGTTT CTGATGCTGA 14580
 TGAGCGACAG CCCACAGCTG CGTGAGTTTC TGGTGCACAA TATCGACAAC ATCGCCAACG 14640
 ATACAGAAGC CTTTATAAAC CGCTACGACC TCAACCGCA TATGATTTCAC AATACTCTGC 14700
 TGATGGTGGAA GGGTAAGCAG CTTGATCGGT TGAAACAAACG TAGCGAGAAA GTCTTGGCGC 14760
 ATCCCCCCCC TAGCAAATGG CTGCAAAAGC GGTTGTACGA TTACCGCTTC TTCCCTCGCTT 14820
 TCGCCGAACA GGATGCCGAG GCAATGAAAG CCGCCTTAGA GCGCGTTTC GATAAAAAAA 14880

78

CCGCGCGTAT GGCTGCCAAA GAAACATTGT CCTATTTCGA TTTCTACCTG CAGCCGAAA	14940
TCGTTACCTA CCCCCAAAATC CCATCCATGC ACGGTTTCGA TTTGGGCATA GATCAAGAAA	15000
TCTCACCGAG GGATTTGATT GTTACGATC CGCTGCCCGC AGACGAATAT CAAGACATCT	15060
TCGATTTAT GAAACAGTAT GACTTGTCTT ACCCGTATGA ATATCTGCAG GATTGGATAG	15120
ATTACTATAC GTTCAAAACC GATAAGCTGG TATTTGGTAA CGCGAAGCGA GAGTGAGCCG	15180
TAAAAACTCTG AGCTCCTGTT TTATAGATTA CAACTTTAGG CCGTCTTAAA GCTGAAAGAT	15240
TTTCGAAAGC TATAAATTGA AGCCCTTCCA CAGTACATAG ATCTGTGTTG TGGCGGGCT	15300
TTACACGCT GATTGCCGGA GAAGAACTCA ACCTGCTGGC AAAACAAGGC ATGAGATCTT	15360
TGCAATAACA TGAGTTGAGA CTTTGCAAA AAAGCCCTTC CCCGACATCC GAAACCCAAA	15420
CACAGGATT CGGCTGTTT CGTACCAAAT ACCTCCTAAT TTTACCCAAA TACCCCTTA	15480
ATCCTCCTCG GACACCCGAT AATCAGGCAT CCGGGCTGCC TTTTAGGCAG CAGCGGGCGC	15540
ATTTAGCCTG TTGGCCGCTT TCAACAGGTT CAAACACATC GCCTTCAGGT GGCTTTGCGC	15600
ACTCACTTG TCATTTCCAA	15620

(2) INFORMATIONS POUR LA SEQ ID NO: 37:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 580 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: Protein
- (B) EMPLACEMENT: 1..580

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 37:

Met	Lys	Phe	Phe	Pro	Ala	Pro	Cys	Leu	Leu	Val	Ile	Leu	Ala	Val	Ile
1															

79

Pro Leu Lys Thr Leu Ala Ala Asp Glu Asn Asp Ala Glu Leu Ile Arg
20 25 30

Ser Met Gln Arg Gln Gln His Ile Asp Ala Glu Leu Leu Thr Asp Ala
35 40 45

Asn Val Arg Phe Glu Gin Pro Leu Glu Lys Asn Asn Tyr Val Leu Ser
50 55 60

Glu Asp Glu Thr Pro Cys Thr Arg Val Asn Tyr Ile Ser Leu Asp Asp
65 70 75 80

Lys Thr Ala Arg Lys Phe Ser Phe Leu Pro Ser Val Leu Met Lys Glu
85 90 95

Thr Ala Phe Lys Thr Gly Met Cys Leu Gly Ser Asn Asn Leu Ser Arg
100 105 110

Leu Gln Lys Ala Ala Gin Gln Ile Leu Ile Val Arg Gly Tyr Leu Thr
115 120 125

Ser Gln Ala Ile Ile Gln Pro Gln Asn Met Asp Ser Gly Ile Leu Lys
130 135 140

Leu Arg Val Ser Ala Gly Glu Ile Gly Asp Ile Arg Tyr Glu Glu Lys
145 150 155 160

Arg Asp Gly Lys Ser Ala Glu Gly Ser Ile Ser Ala Phe Asn Asn Lys
165 170 175

Phe Pro Leu Tyr Arg Asn Lys Ile Leu Asn Leu Arg Asp Val Glu Gln
180 185 190

Gly Leu Glu Asn Leu Arg Arg Leu Pro Ser Val Lys Thr Asp Ile Gln
195 200 205

Ile Ile Pro Ser Glu Glu Gly Lys Ser Asp Leu Gln Ile Lys Trp
210 215 220

Gln Gln Asn Lys Pro Ile Arg Phe Ser Ile Gly Ile Asp Asp Ala Gly
225 230 235 240

Gly Lys Thr Thr Gly Lys Tyr Gln Gly Asn Val Ala Leu Ser Phe Asp
245 250 255

80

Asn Pro Leu Gly Leu Ser Asp Leu Phe Tyr Val Ser Tyr Gly Arg Gly
 260 265 270

Leu Val His Lys Thr Asp Leu Thr Asp Ala Thr Gly Thr Glu Thr Glu
 275 280 285

Ser Gly Ser Arg Ser Tyr Ser Val His Tyr Ser Val Pro Val Lys Lys
 290 295 300

Trp Leu Phe Ser Phe Asn His Asn Gly His Arg Tyr His Glu Ala Thr
 305 310 315 320

Glu Gly Tyr Ser Val Asn Tyr Asp Tyr Asn Gly Lys Gln Tyr Gln Ser
 325 330 335

Ser Leu Ala Ala Glu Arg Met Leu Trp Arg Asn Arg Phe His Lys Thr
 340 345 350

Ser Val Gly Met Lys Leu Trp Thr Arg Gln Thr Tyr Lys Tyr Ile Asp
 355 360 365

Asp Ala Glu Ile Glu Val Gln Arg Arg Arg Ser Ala Gly Trp Glu Ala
 370 375 380

Glu Leu Arg His Arg Ala Tyr Leu Asn Arg Trp Gln Leu Asp Gly Lys
 385 390 395 400

Leu Ser Tyr Lys Arg Gly Thr Gly Met Arg Gln Ser Met Pro Ala Pro
 405 410 415

Glu Glu Asn Gly Gly Thr Ile Pro Gly Thr Ser Arg Met Lys Ile
 420 425 430

Ile Thr Ala Gly Leu Asp Ala Ala Ala Pro Phe Met Leu Gly Lys Gln
 435 440 445

Gln Phe Phe Tyr Ala Thr Ala Ile Gln Ala Gln Trp Asn Lys Thr Pro
 450 455 460

Leu Val Ala Gln Asp Lys Leu Ser Ile Gly Ser Arg Tyr Thr Val Arg
 465 470 475 480

Gly Phe Asp Gly Glu Gln Ser Leu Phe Gly Glu Arg Gly Phe Tyr Trp
 485 490 495

81

Gln Asn Thr L^{eu} Thr Trp Tyr Phe His Pro Asn His Gln Phe Tyr L^{eu}
500 505 510

Gly Ala Asp Tyr Gly Arg Val Ser Gly Glu Ser Ala Gln Tyr Val Ser
515 520 525

Gly Lys Gln L^{eu} Met Gly Ala Val Val Gly Phe Arg Gly Gly His Lys
530 535 540

Val Gly Gly Met Phe Ala Tyr Asp L^{eu} Phe Ala Gly Lys Pro L^{eu} His
545 550 555 560

Lys Pro Lys Gly Phe Gln Thr Thr Asn Thr Val Tyr Gly Phe Asn L^{eu}
565 570 575

Asn Tyr Ser Phe
580

(2) INFORMATIONS POUR LA SEQ ID NO: 38:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1981 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(ix) CARACTERISTIQUE:

- (A) NOM/CLÉ: Peptide
- (B) EMPLACEMENT: 1...1981

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 38:

Met Asn Lys Gly L^{eu} His Arg Ile Ile Phe Ser Lys Lys His Ser Thr
1 5 10 15

Met Val Ala Val Ala Glu Thr Ala Asn Ser Gln Gly Lys Gly Lys Gln
20 25 30

Ala Gly Ser Ser Val Ser Val Ser L^{eu} Lys Thr Ser Gly Asp L^{eu} Cys
35 40 45

82

Gly Lys Leu Lys Thr Thr Leu Lys Thr Leu Val Cys Ser Leu Val Ser
 50 55 60

Leu Ser Met Val Leu Pro Ala His Ala Gln Ile Thr Thr Asp Lys Ser
 65 70 75 80

Ala Pro Lys Asn Gln Gln Val Val Ile Leu Lys Thr Asn Thr Gly Ala
 85 90 95

Pro Leu Val Asn Ile Gln Thr Pro Asn Gly Arg Gly Leu Ser His Asn
 100 105 110

Arg Tyr Thr Gln Phe Asp Val Asp Asn Lys Gly Ala Val Leu Asn Asn
 115 120 125

Asp Arg Asn Asn Asn Pro Phe Leu Val Lys Gly Ser Ala Gln Leu Ile
 130 135 140

Leu Asn Glu Val Arg Gly Thr Ala Ser Lys Leu Asn Gly Ile Val Thr
 145 150 155 160

Val Gly Gly Gln Lys Ala Asp Val Ile Ile Ala Asn Pro Asn Gly Ile
 165 170 175

Thr Val Asn Gly Gly Phe Lys Asn Val Gly Arg Gly Ile Leu Thr
 180 185 190

Ile Gly Ala Pro Gln Ile Gly Lys Asp Gly Ala Leu Thr Gly Phe Asp
 195 200 205

Val Arg Gln Gly Thr Leu Thr Val Gly Ala Ala Gly Trp Asn Asp Lys
 210 215 220

Gly Gly Ala Asp Tyr Thr Gly Val Leu Ala Arg Ala Val Ala Leu Gln
 225 230 235 240

Gly Lys Leu Gln Gly Lys Asn Leu Ala Val Ser Thr Gly Pro Gln Lys
 245 250 255

Val Asp Tyr Ala Ser Gly Glu Ile Ser Ala Gly Thr Ala Ala Gly Thr
 260 265 270

Lys Pro Thr Ile Ala Leu Asp Thr Ala Ala Leu Gly Gly Met Tyr Ala
 275 280 285

Asp Ser Ile Thr Leu Ile Ala Asn Glu Lys Gly Val Gly Val Lys Asn
290 295 300

Ala Gly Thr Leu Glu Ala Ala Lys Gln Leu Ile Val Thr Ser Ser Gly
305 310 315 320

Arg Ile Glu Asn Ser Gly Arg Ile Ala Thr Thr Ala Asp Gly Thr Glu
325 330 335

Ala Ser Pro Thr Tyr Leu Ser Ile Glu Thr Thr Glu Lys Gly Ala Ala
340 345 350

Gly Thr Phe Ile Ser Asn Gly Gly Arg Ile Glu Ser Lys Gly Leu Leu
355 360 365

Val Ile Glu Thr Gly Glu Asp Ile Ser Leu Arg Asn Gly Ala Val Val
370 375 380

Gln Asn Asn Gly Ser Arg Pro Ala Thr Thr Val Leu Asn Ala Gly His
385 390 395 400

Asn Leu Val Ile Glu Ser Lys Thr Asn Val Asn Asn Ala Lys Gly Ser
405 410 415

Ala Asn Leu Ser Ala Gly Gly Arg Thr Thr Ile Asn Asp Ala Thr Ile
420 425 430

Gln Ala Gly Ser Ser Val Tyr Ser Ser Thr Lys Gly Asp Thr Glu Leu
435 440 445

Gly Glu Asn Thr Arg Ile Ile Ala Glu Asn Val Thr Val Leu Ser Asn
450 455 460

Gly Ser Ile Gly Ser Ala Ala Val Ile Glu Ala Lys Asp Thr Ala His
465 470 475 480

Ile Glu Ser Gly Lys Pro Leu Ser Leu Glu Thr Ser Thr Val Ala Ser
485 490 495

Asn Ile Arg Leu Asn Asn Gly Asn Ile Lys Gly Gly Lys Gln Leu Ala
500 505 510

Leu Leu Ala Asp Asp Asn Ile Thr Ala Lys Thr Thr Asn Leu Asn Thr
515 520 525

Pro Gly Asn Leu Tyr Val His Thr Gly Lys Asp Leu Asn Leu Asn Val
530 535 540

Asp Lys Asp Leu Ser Ala Ala Ser Ile His Leu Lys Ser Asp Asn Ala
545 550 555 560

Ala His Ile Thr Gly Thr Ser Lys Thr Leu Thr Ala Ser Lys Asp Met
565 570 575

Gly Val Glu Ala Gly Leu Leu Asn Val Thr Asn Thr Asn Leu Arg Thr
580 585 590

Asn Ser Gly Asn Leu His Ile Gln Ala Ala Lys Gly Asn Ile Gln Leu
595 600 605

Arg Asn Thr Lys Leu Asn Ala Ala Lys Ala Leu Glu Thr Thr Ala Leu
610 615 620

Gln Gly Asn Ile Val Ser Asp Gly Leu His Ala Val Ser Ala Asp Gly
625 630 635 640

His Val Ser Leu Leu Ala Asn Gly Asn Ala Asp Phe Thr Gly His Asn
645 650 655

Thr Leu Thr Ala Lys Ala Asp Val Asn Ala Gly Ser Val Gly Lys Gly
660 665 670

Arg Leu Lys Ala Asp Asn Thr Asn Ile Thr Ser Ser Gly Asp Ile
675 680 685

Thr Leu Val Ala Gly Asn Gly Ile Gln Leu Gly Asp Gly Lys Gln Arg
690 695 700

Asn Ser Ile Asn Gly Lys His Ile Ser Ile Lys Asn Asn Gly Gly Asn
705 710 715 720

Ala Asp Leu Lys Asn Leu Asn Val His Ala Lys Ser Gly Ala Leu Asn
725 730 735

Ile His Ser Asp Arg Ala Leu Ser Ile Glu Asn Thr Lys Leu Glu Ser
740 745 750

Thr His Asn Thr His Leu Asn Ala Gln His Glu Arg Val Thr Leu Asn
755 760 765

85

Gln Val Asp Ala Tyr Ala His Arg His Leu Ser Ile Thr Gly Ser Gln
770 775 780

Ile Trp Gln Asn Asp Lys Leu Pro Ser Ala Asn Lys Leu Val Ala Asn
785 790 795 800

Gly Val Leu Ala Leu Asn Ala Arg Tyr Ser Gln Ile Ala Asp Asn Thr
805 810 815

Thr Leu Arg Ala Gly Ala Ile Asn Leu Thr Ala Gly Thr Ala Leu Val
820 825 830

Lys Arg Gly Asn Ile Asn Trp Ser Thr Val Ser Thr Lys Thr Leu Glu
835 840 845

Asp Asn Ala Glu Leu Lys Pro Leu Ala Gly Arg Leu Asn Ile Glu Ala
850 855 860

Gly Ser Gly Thr Leu Thr Ile Glu Pro Ala Asn Arg Ile Ser Ala His
865 870 875 880

Thr Asp Leu Ser Ile Lys Thr Gly Gly Lys Leu Leu Leu Ser Ala Lys
885 890 895

Gly Gly Asn Ala Gly Ala Pro Ser Ala Gln Val Ser Ser Leu Glu Ala
900 905 910

Lys Gly Asn Ile Arg Leu Val Thr Gly Glu Thr Asp Leu Arg Gly Ser
915 920 925

Lys Ile Thr Ala Gly Lys Asn Leu Val Val Ala Thr Thr Lys Gly Lys
930 935 940

Leu Asn Ile Glu Ala Val Asn Asn Ser Phe Ser Asn Tyr Phe Pro Thr
945 950 955 960

Gln Lys Ala Ala Glu Leu Asn Gln Lys Ser Lys Glu Leu Glu Gln Gln
965 970 975

Ile Ala Gln Leu Lys Lys Ser Ser Pro Lys Ser Lys Leu Ile Pro Thr
980 985 990

Leu Gln Glu Glu Arg Asp Arg Leu Ala Phe Tyr Ile Gln Ala Ile Asn
995 1000 1005

86

Lys Glu Val Lys Gly Lys Lys Pro Lys Gly Lys Glu Tyr Leu Gln Ala
 1010 1015 1020

Lys Leu Ser Ala Gln Asn Ile Asp Leu Ile Ser Ala Gln Gly Ile Glu
 1025 1030 1035 1040

Ile Ser Gly Ser Asp Ile Thr Ala Ser Lys Lys Leu Asn Leu His Ala
 1045 1050 1055

Ala Gly Val Leu Pro Lys Ala Ala Asp Ser Glu Ala Ala Ala Ile Leu
 1060 1065 1070

Ile Asp Gly Ile Thr Asp Gln Tyr Glu Ile Gly Lys Pro Thr Tyr Lys
 1075 1080 1085

Ser His Tyr Asp Lys Ala Ala Leu Asn Lys Pro Ser Arg Leu Thr Gly
 1090 1095 1100

Arg Thr Gly Val Ser Ile His Ala Ala Ala Leu Asp Asp Ala Arg
 1105 1110 1115 1120

Ile Ile Ile Gly Ala Ser Glu Ile Lys Ala Pro Ser Gly Ser Ile Asp
 1125 1130 1135

Ile Lys Ala His Ser Asp Ile Val Leu Glu Ala Gly Gln Asn Asp Ala
 1140 1145 1150

Tyr Thr Phe Leu Lys Thr Lys Gly Lys Ser Gly Lys Ile Ile Arg Lys
 1155 1160 1165

Thr Lys Phe Thr Ser Thr Arg Asp His Leu Ile Met Pro Ala Pro Val
 1170 1175 1180

Glu Leu Thr Ala Asn Gly Ile Thr Leu Gln Ala Gly Gly Asn Ile Glu
 1185 1190 1195 1200

Ala Asn Thr Thr Arg Phe Asn Ala Pro Ala Gly Lys Val Thr Leu Val
 1205 1210 1215

Ala Gly Glu Glu Leu Gln Leu Ala Glu Glu Gly Ile His Lys His
 1220 1225 1230

Glu Leu Asp Val Gln Lys Ser Arg Arg Phe Ile Gly Ile Lys Val Gly
 1235 1240 1245

Lys Ser Asn Tyr Ser Lys Asn Glu Leu Asn Glu Thr Lys Leu Pro Val
 1250 1255 1260

Arg Val Val Ala Gln Thr Ala Ala Thr Arg Ser Gly Trp Asp Thr Val
 1265 1270 1275 1280

Leu Glu Gly Thr Glu Phe Lys Thr Thr Leu Ala Gly Ala Asp Ile Gln
 1285 1290 1295

Ala Gly Val Gly Glu Lys Ala Arg Val Asp Ala Lys Ile Ile Leu Lys
 1300 1305 1310

Gly Ile Val Asn Arg Ile Gln Ser Glu Glu Lys Leu Glu Thr Asn Ser
 1315 1320 1325

Thr Val Trp Gln Lys Gln Ala Gly Arg Gly Ser Thr Ile Glu Thr Leu
 1330 1335 1340

Lys Leu Pro Ser Phe Glu Ser Pro Thr Pro Pro Lys Leu Ser Ala Pro
 1345 1350 1355 1360

Gly Gly Tyr Ile Val Asp Ile Pro Lys Gly Asn Leu Lys Thr Glu Ile
 1365 1370 1375

Glu Lys Leu Ser Lys Gln Pro Glu Tyr Ala Tyr Leu Lys Gln Leu Gln
 1380 1385 1390

Val Ala Lys Asn Ile Asn Trp Asn Gln Val Gln Leu Ala Tyr Asp Arg
 1395 1400 1405

Trp Asp Tyr Lys Gln Glu Gly Leu Thr Glu Ala Gly Ala Ala Ile Ile
 1410 1415 1420

Ala Leu Ala Val Thr Val Val Thr Ser Gly Ala Gly Thr Gly Ala Val
 1425 1430 1435 1440

Leu Gly Leu Asn Gly Ala Ala Ala Ala Ala Thr Asp Ala Ala Phe Ala
 1445 1450 1455

Ser Leu Ala Ser Gln Ala Ser Val Ser Phe Ile Asn Asn Lys Gly Asp
 1460 1465 1470

Val Gly Lys Thr Leu Lys Glu Leu Gly Arg Ser Ser Thr Val Lys Asn
 1475 1480 1485

88

L^eu V^al V^al A^la A^la A^la T^hr A^la G^ly V^al A^la A^sp L^ys I^le G^ly A^la
 1490 1495 1500

S^er A^la L^eu A^sn A^sn V^al S^er A^sp L^ys G^ln T^rp I^le A^sn A^sn L^eu T^hr
 1505 1510 1515 1520

V^al A^sn L^eu A^la A^sn A^la G^ly S^er A^la A^la L^eu I^le A^sn T^hr A^la V^al
 1525 1530 1535

A^sn G^ly G^ly S^er L^eu L^ys A^sp A^sn L^eu G^lu A^la A^sn I^le L^eu A^la A^la
 1540 1545 1550

L^eu V^al A^sn T^hr A^la Hⁱs G^ly G^lu A^la A^la S^er L^ys I^le L^ys G^ln L^eu
 1555 1560 1565

A^sp G^ln Hⁱs T^yr I^le V^al Hⁱs L^ys I^le A^la Hⁱs A^la I^le A^la G^ly C^ys
 1570 1575 1580

A^la A^la A^la A^la A^la A^sn L^ys G^ly L^ys C^ys G^ln A^sp G^ly A^la I^le G^ly
 1585 1590 1595 1600

A^la A^la V^al G^ly G^lu I^le V^al G^ly G^lu A^la L^eu T^hr A^sn G^ly L^ys A^sn
 1605 1610 1615

P^rc A^sp T^hr L^eu T^hr A^la L^ys G^lu A^rg G^lu G^ln I^le L^eu A^la T^yr S^er
 1620 1625 1630

L^ys L^eu V^al A^la G^ly T^hr V^al S^er G^ly V^al V^al G^ly A^sp V^al A^sn
 1635 1640 1645

A^la A^la A^la A^sn A^la A^la G^lu V^al A^la V^al L^ys A^sn A^sn G^ln L^eu S^er
 1650 1655 1660

A^sp L^ys G^lu G^ly A^rg G^lu P^he A^sp A^sn G^lu M^et T^hr A^la C^ys A^la L^ys
 1665 1670 1675 1680

G^ln A^sn A^sn P^ro G^ln L^eu C^ys A^rg L^ys A^sn T^hr V^al L^ys L^ys T^yr G^ln
 1685 1690 1695

A^sn V^al A^la A^sp L^ys A^rg L^eu A^la A^la S^er I^le A^la I^le C^ys T^hr A^sp
 1700 1705 1710

I^le S^er A^rg S^er T^hr G^lu C^ys A^rg T^hr I^le A^rg L^ys G^ln Hⁱs L^eu I^le
 1715 1720 1725